US008580556B2

(12) **United States Patent**
McBride et al.

(10) **Patent No.: US 8,580,556 B2**
(45) **Date of Patent: Nov. 12, 2013**

(54) **SULFONYLUREA-RESPONSIVE REPRESSOR PROTEINS**

(75) Inventors: **Kevin E McBride**, Davis, CA (US); **Michael W Lassner**, Urbandale, IA (US); **Loren L Looger**, Madison, AL (US); **Brian McGonigle**, Wilmington, DE (US)

(73) Assignees: **E. I. du Pont de Nemours and Company**, Wilmington, DE (US); **Pioneer Hi-Bred International Inc.**, Johnston, IA (US)

(*) Notice: Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 0 days.

(21) Appl. No.: **13/563,854**

(22) Filed: **Aug. 1, 2012**

(65) **Prior Publication Data**

US 2012/0295354 A1 Nov. 22, 2012

Related U.S. Application Data

(63) Continuation of application No. 12/603,739, filed on Oct. 22, 2009, now Pat. No. 8,257,956.

(60) Provisional application No. 61/108,917, filed on Oct. 28, 2008.

(51) **Int. Cl.**
C12N 1/20 (2006.01)
C07H 21/04 (2006.01)

(52) **U.S. Cl.**
USPC **435/252.3**; 536/23.7; 435/252.33; 435/252.8

(58) **Field of Classification Search**
None
See application file for complete search history.

(56) **References Cited**

U.S. PATENT DOCUMENTS

5,464,758	A	11/1995	Gossen et al.	
5,589,362	A	12/1996	Bujard et al.	
5,605,011	A *	2/1997	Bedbrook et al.	47/58.1 R
5,650,298	A	7/1997	Bujard et al.	
5,654,168	A	8/1997	Bujard et al.	
5,723,765	A	3/1998	Oliver et al.	
5,789,156	A	8/1998	Bujard et al.	
5,814,618	A	9/1998	Bujard et al.	
5,851,796	A	12/1998	Schatz	
5,859,310	A	1/1999	Bujard et al.	
5,866,755	A	2/1999	Bujard et al.	
5,888,981	A	3/1999	Bujard et al.	
5,912,411	A	6/1999	Bujard et al.	
5,917,122	A	6/1999	Byrne	
5,922,927	A	7/1999	Bujard et al.	
5,925,808	A	7/1999	Oliver et al.	
5,977,441	A	11/1999	Oliver et al.	
6,004,941	A	12/1999	Bujard et al.	
6,087,166	A	7/2000	Baron et al.	
6,136,954	A	10/2000	Bujard et al.	
6,242,667	B1	6/2001	Bujard et al.	
6,252,136	B1	6/2001	Bujard et al.	
6,271,341	B1	8/2001	Baron et al.	
6,271,348	B1	8/2001	Bujard et al.	
6,783,756	B2	8/2004	Bujard et al.	
6,914,124	B2	7/2005	Bujard et al.	
2003/0186281	A1	10/2003	Hillen	
2003/0208783	A1	11/2003	Hillen et al.	
2004/0133944	A1	7/2004	Hake et al.	
2004/0148649	A1	7/2004	Lydiate et al.	
2005/0034187	A1	2/2005	Golovko et al.	
2005/0042752	A1	2/2005	Crete et al.	
2005/0216976	A1	9/2005	Meagher et al.	

FOREIGN PATENT DOCUMENTS

CN	1385527	12/2002
EP	0 494 724 A2	7/1992
WO	96/04393	2/1996
WO	96/40892	12/1996
WO	96/40946	12/1996
WO	99/27108	6/1999
WO	00/75347	12/2000
WO	03/029453	4/2003
WO	03/056021	7/2003
WO	2004/020645	3/2004
WO	2005/123923	12/2005
WO	2006/124502	11/2006
WO	2007/133797	11/2007
WO	2010/062518	6/2010

OTHER PUBLICATIONS

Scholz, et al., Activity reversal of Tet repressor caused by single amino acid exchanges, Mol. Microbiology, (2004), 53(3): 777-789.

Wharton, et al., Substituting an a-Helix Switches the Sequence-Specific DNA Interactions of a Repressor, Cell, (1984), 38:361-369.

Dallmann, et al., The isolated N-terminal DNA binding domain of the c repressor of bacteriophage 16-3 is functional in DNA binding in vivo and in vitro, Mol Gen Genet (1991), 227:106-112.

Pog Chapter 15, Advanced transgenic technology, (2005), 15:299-320.

Kramer, et al., An engineered epigenetic transgene switch in mammalian cells, Nature Biotechnology, (2004), 22 (7):867-870.

Wharton, et al., Changing the binding specificity of a repressor by redesigning an a-helix, Nature, (1985), 316:601-605.

(Continued)

Primary Examiner — Padma V Baskar

(57) **ABSTRACT**

Compositions and methods relating to the use of sulfonylurea-responsive repressors are provided. Compositions include polypeptides that specifically bind to an operator, wherein the specific binding is regulated by a sulfonylurea compound. Compositions also include polynucleotides encoding the polypeptides as well as constructs, vectors, prokaryotic and eukaryotic cells, and eukaryotic organisms including plants and seeds comprising the polynucleotide, and/or produced by the methods. Also provided are methods to provide a sulfonylurea-responsive repressor to a cell or organism, and to regulate expression of a polynucleotide of interest in a cell or organism, including a plant or plant cell.

2 Claims, 11 Drawing Sheets

(56) **References Cited**

OTHER PUBLICATIONS

Hollis, et al., A repressor heterodimer binds to a chimeric operator, Proc. Natl. Acad. Sci. USA (1988), 85:5834-5838.

Degenkolb, et al., Structural Requirements of Tetracycline-Tet Repressor Interaction: Determination of Equilibrium Binding Constants for Tetracycline Analogs with the Tet Repressor, Antimicrobial Agents and Chemotherapy, (1991), 35(8):1591-1595.

Bustos, et al., Functional domains of the AraC protein, Proc. Natl. Acad. Sci. USA (1993), 90:5638-5642.

Lallo, et al., A two-hybrid system based on chimeric operator recognition for studying protein homo/heterodimerization in *Escherichia coli*, Microbiology, (2001), 147:1651-1656.

Backes, et al., Combinations of the a-Helix-turn-a-Helix Motif of TetR with Respective Residues from LacI or 434Cro: DNA Recognition, Inducer Binding, and Urea-Dependent Denaturation, Biochemistry (1997), 36:5311-5322.

Wilcken-Bergmann, et al., 72 Residues of gal repressor fused to B-galactosidase repress the gal operon of *E. coli*, EMBO Journal (1983), 2(8):1271-1274.

Zeng, et al., Detection of tetramerization domains in vivo by cooperative DNA binding to tandem n operator sites, Gene (1997), 185:245-249.

Helbl, et al., Stepwise Selection of TetR Variants Recognizing tet Operator 4C with High Affinity and Specificity, J. Mol. Biol., (1998), 276:313-318.

Helbl, et al., Stepwise Selection of TetR Variants Recognizing tet Operator 6C with High Affinity and Specificity, J. Mol. Biol. (1998), 276:319-324.

Muller, et al., Characterization of non-inducible Tet repressor mutants suggests conformational changes necessary for induction, Nature Structural Biology, (1995), 2(8):693-703.

Wang, et al., Transcriptional Repression in Saccharomyces cerevisiae by a SIN3-LexA Fusion Protein, Mol. and Cell. Biology, (1993), 13(3):1805-1814.

Isackson, et al., Dominant negative mutations in the Tn10 tet repressor: Evidence for use of the conserved helix-turn-helix motif in DNA binding, Proc. Natl. Acad. Sci. USA, (1985), 82:6226-6230.

Bertram, et al., Phenotypes of Combined Tet Repressor Mutants for Effector and Operator recognition and Allostery, J. Mol. Microbiology and Biotechnol., (2004), 8:104-110.

Urlinger, et al., Exploring the sequence space for tetracycline-dependent transcriptional activators: Novel mutations yield expanded range and sensitivity, PNAS (2000), 97(14):7963-7968.

Berens, et al., Gene regulation by tetracyclines, Eur. J. Biochem., (2003), 270:3109-3121.

Kamionka, et al., Two mutations in the tetracycline repressor change the inducer anhydrotetracycline to a corepressor, Nucleic Acids Research, (2004), 32(2):842-847.

Scholz, et al., Teaching TetR to Recognize a New Inducer, J. Mol. Biol., (2003), 329:217-227.

Ramos, et al., The TetR Family of Transcriptional Repressors, Microbiology and Mol. Biology Rev., (2005), 69 (2):326-356.

Orth, et al., Structural basis of gene regulation by the tetracycline inducible Tet repressor—operator system, Nature Structural Biology, (2000), 7(3):215-219.

Patel, et al., Phenobarbital and sulfonylurea-inducible operons encoding herbicide metabolizing cytochromes P-450 in *Streptomyces griseolus*, Gene, Mar. 1, 1992, vol. 112(1) pp. 67-76 (abstract only).

De Veylder Lieven et al., Herbicide Safener-Inducible Gene Expression in *Arabidopsis thaliana*,Plant and Cell Physiology, 1997, pp. 568-577, vol. 38(5).

Gatz et al., Stringent Repression and Homogenous De-repression by Tetracycline of a Modified CaMV 35S Promoter in Intact Transgenic Tobacco Plants, Plant Journal, 1992, pp. 397-404, vol. 2(3), Blackwell Scientific Publications, Oxford, GB.

U.S. Appl. No. 13/086,765, filed Apr. 14, 2011.

* cited by examiner

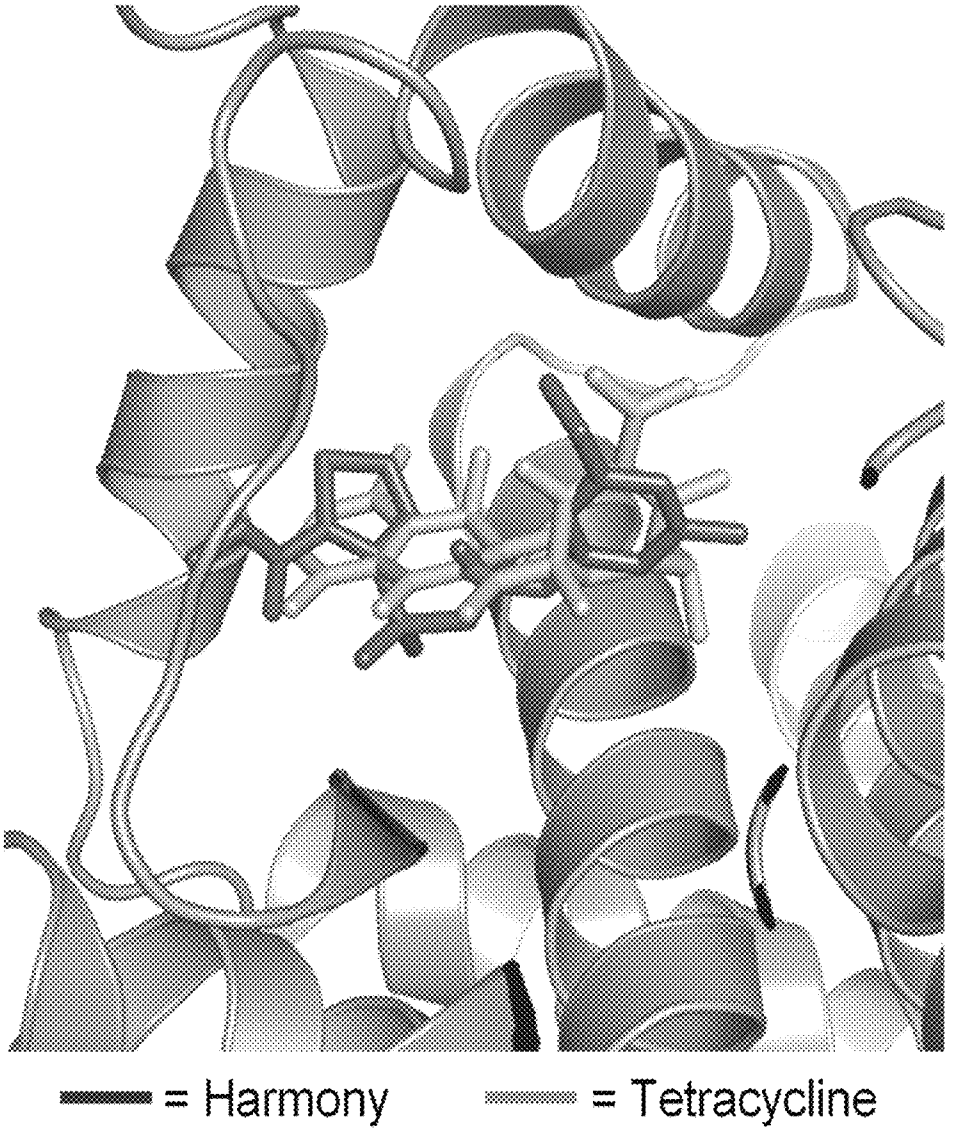


Figure 1

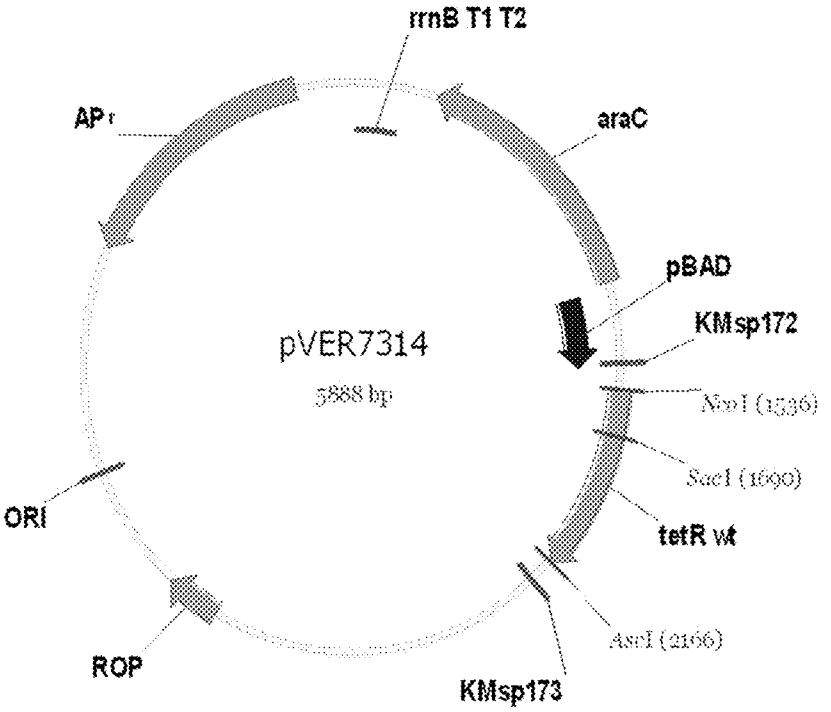

rrnB T1 T2
APr
araC
pBAD
KMsp172
pVER7314
5888 bp
NcoI (1536)
SacI (1690)
tetR wt
AscI (2166)
KMsp173
ROP
ORI

Figure 2

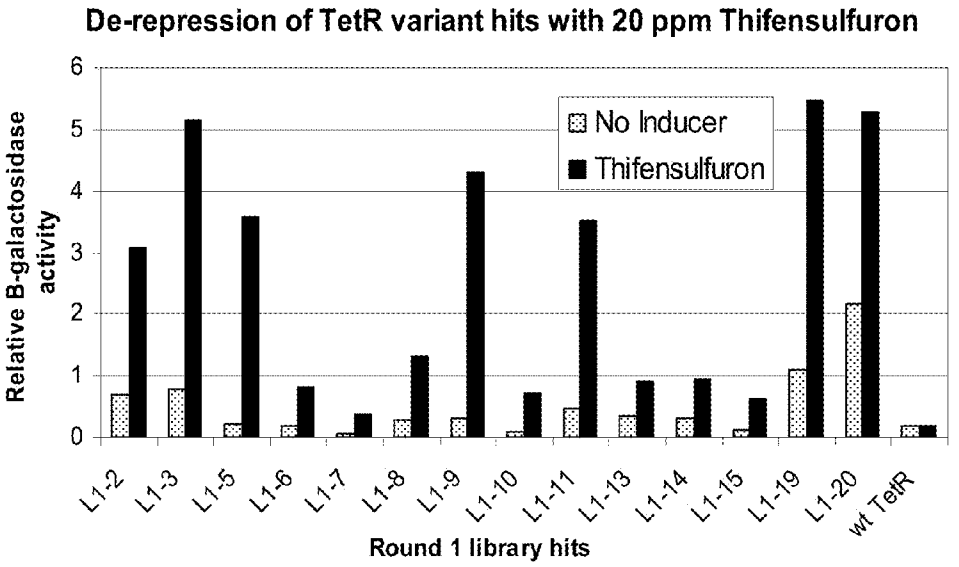

De-repression of TetR variant hits with 20 ppm Thifensulfuron
Relative B-galactosidase activity
0
1
2
3
4
5
6
No Inducer
Thifensulfuron
L1-2
L1-3
L1-5
L1-6
L1-7
L1-8
L1-9
L1-10
L1-11
L1-13
L1-14
L1-15
L1-19
L1-20
wt TetR
Round 1 library hits

Figure 3

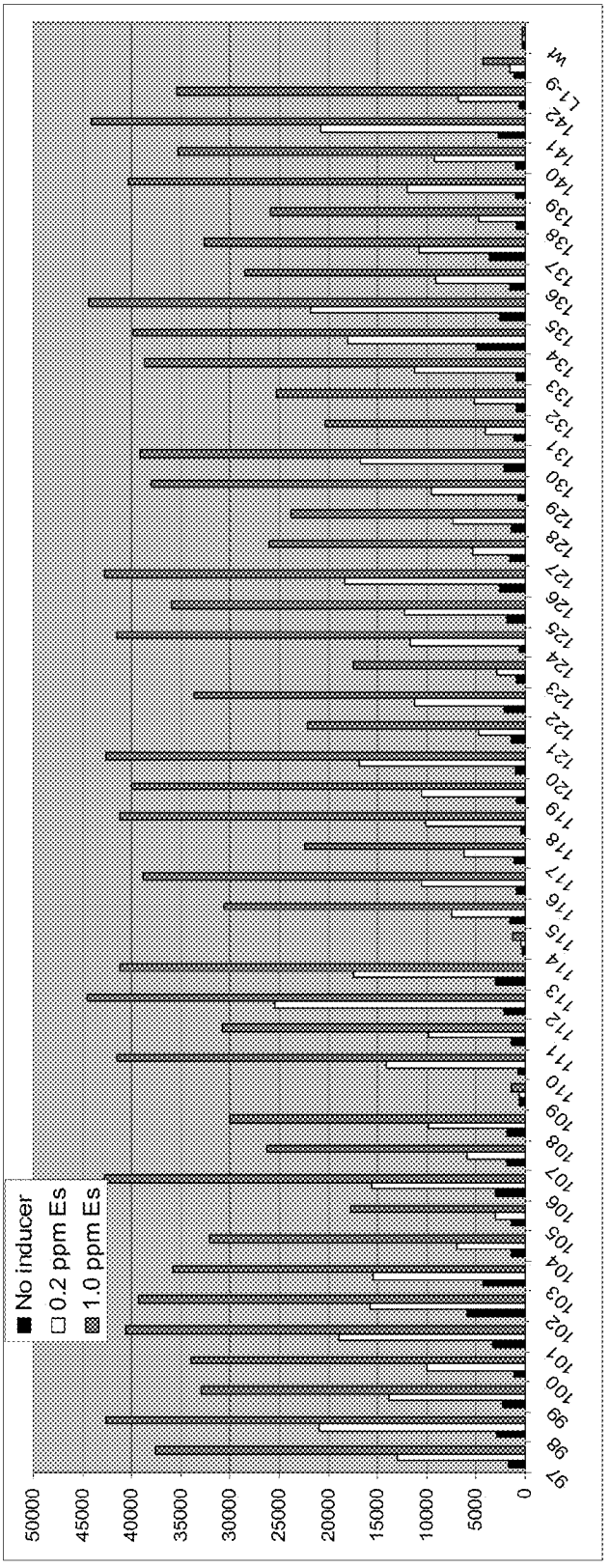

β-galactosidase activity
50000
45000
40000
35000
30000
25000
20000
15000
10000
5000
0
No inducer
0.2 ppm Es
1.0 ppm Es
97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 L1-9 wt
Clone ID

Figure 4

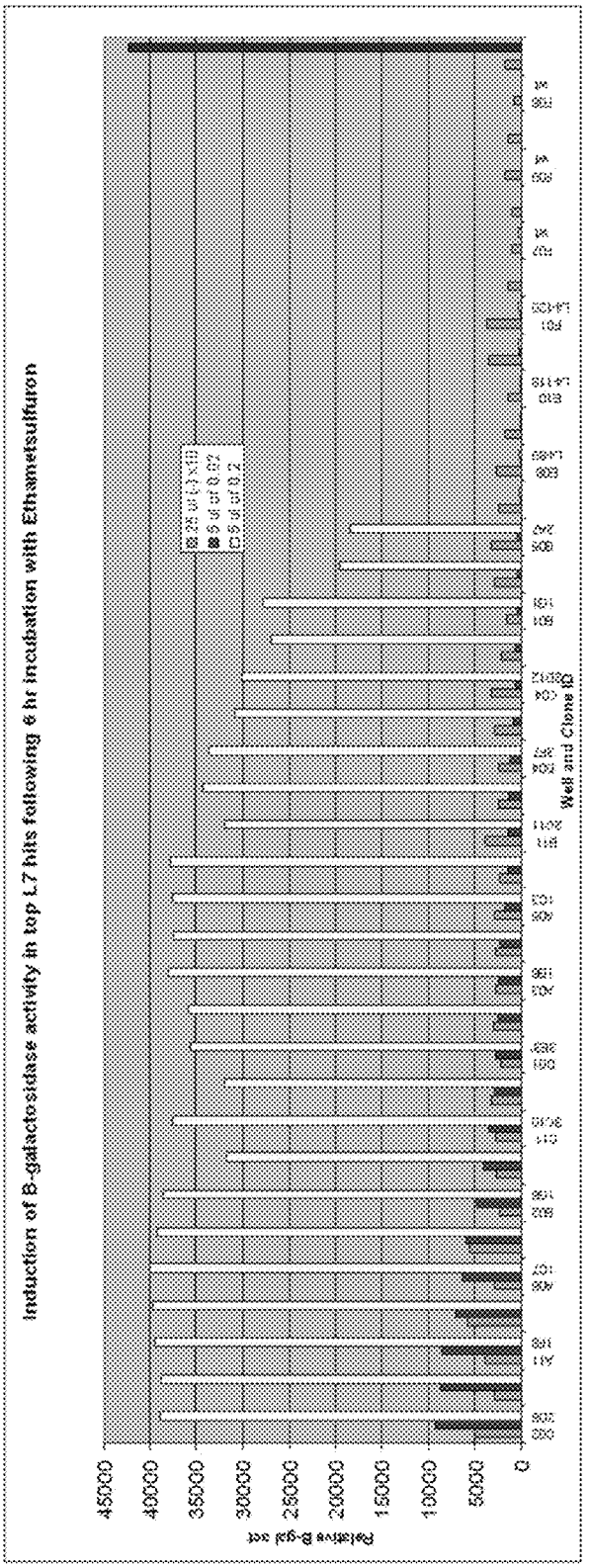

Induction of B-galactosidase activity in top L7 hits following 6 hr incubation with Ethametsulfuron
Relative B-gal act
Well and Clone ID
0
5000
10000
15000
20000
25000
30000
35000
40000
45000

Figure 5

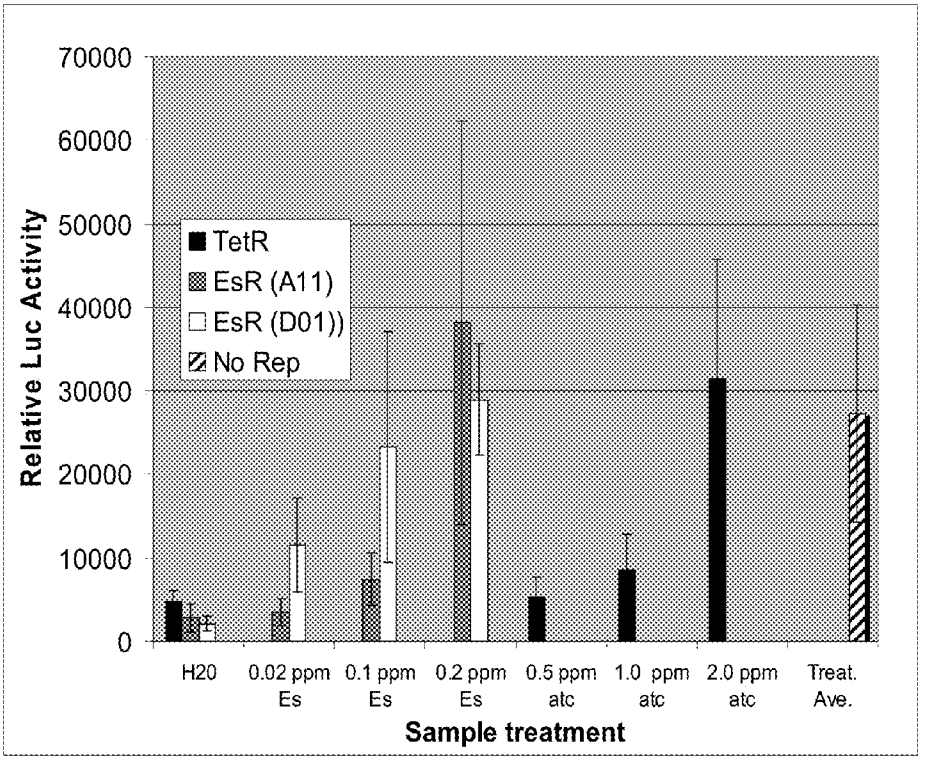
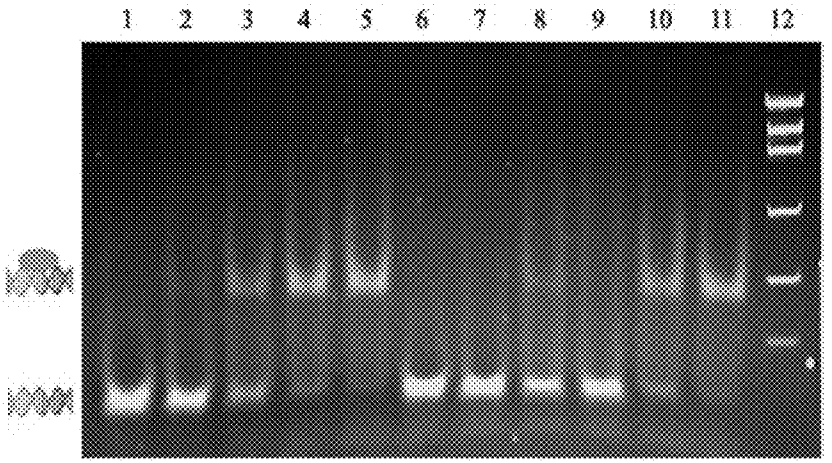

70000
60000
50000
40000
30000
20000
10000
0
Relative Luc Activity
TetR
EsR (A11)
EsR (D01))
No Rep
H20
0.02 ppm Es
0.1 ppm Es
0.2 ppm Es
0.5 ppm atc
1.0 ppm atc
2.0 ppm atc
Treat. Ave.
Sample treatment

Figure 6



TetO	(5 pmol)	+	+	+	+	+	+	–	+	+	+	+
TetO Ctrl	(5 pmol)	–	–	–	–	–	–	+	–	–	–	–
L7A11	(in pmol)	–	5	15	25	35	–	35	35	35	35	35
BSA	(in pmol)	–	–	–	–	–	35	–	–	–	–	–
Es	(in pmol)	–	–	–	–	–	–	–	75	150	–	–
Cs	(in pmol)	–	–	–	–	–	–	–	–	–	150	–
Atc	(in pmol)	–	–	–	–	–	–	–	–	–	–	150

Figure 7

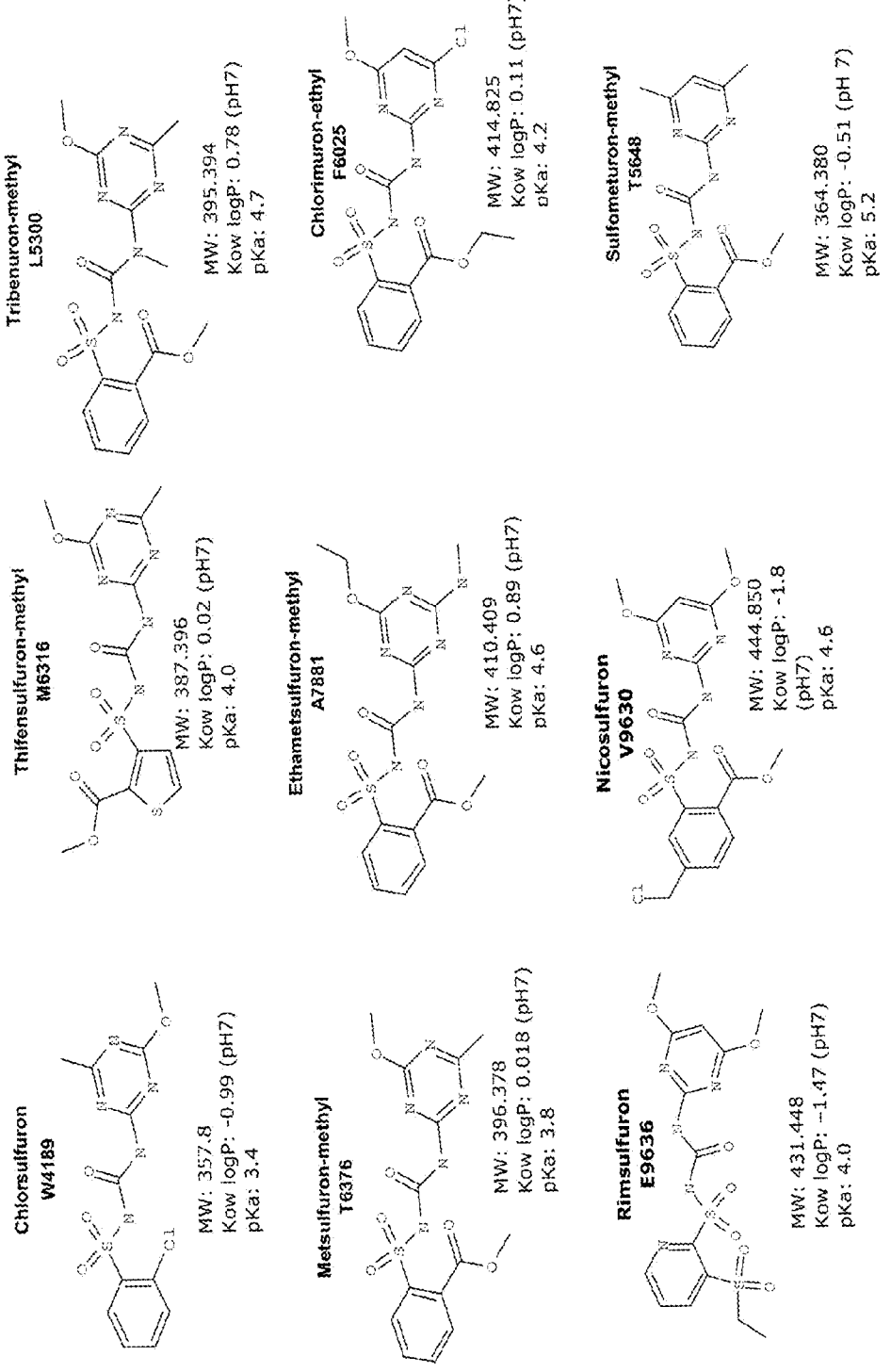
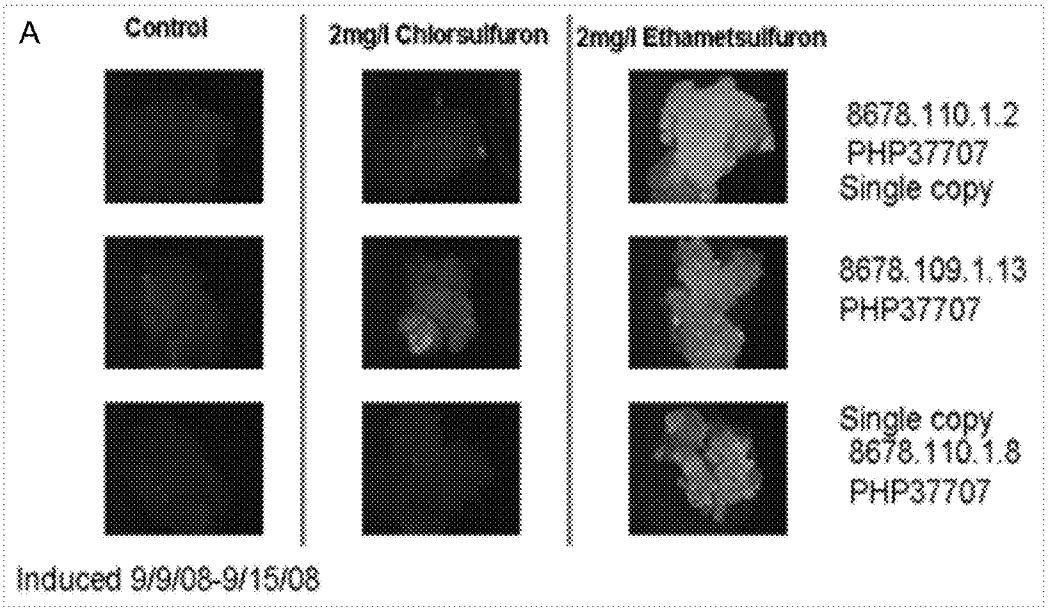

Tribenuron-methyl
L5300
MW: 395.394
Kow logP: 0.78 (pH7)
pKa: 4.7
Chlorimuron-ethyl
F6025
MW: 414.825
Kow logP: 0.11 (pH7)
pKa: 4.2
Sulfometuron-methyl
T5648
MW: 364.380
Kow logP: -0.51 (pH 7)
pKa: 5.2
Thifensulfuron-methyl
M6316
MW: 387.396
Kow logP: 0.02 (pH7)
pKa: 4.0
Ethametsulfuron-methyl
A7881
MW: 410.409
Kow logP: 0.89 (pH7)
pKa: 4.6
Nicosulfuron
V9630
MW: 444.850
Kow logP: -1.8
(pH7)
pKa: 4.6
Chlorsulfuron
W4189
MW: 357.8
Kow logP: -0.99 (pH7)
pKa: 3.4
Metsulfuron-methyl
T6376
MW: 396.378
Kow logP: 0.018 (pH7)
pKa: 3.8
Rimsulfuron
E9636
MW: 431.448
Kow logP: -1.47 (pH7)
pKa: 4.0

Figure 8

		Library					
Backbone position	WT Seq	L1 (planned)	L1 (1st screen Ts. 2nd screen all SU's in table 3)	L2 (Cs)	L4 (Es)	L6 (Cs)	L7 (Es)
49	R	-	-	-	GQSR	-	-
52	L	-	-	-	LM	-	-
55	L	-	-	-	**LM**	-	**M**L
60	L	ALKM	AL[K]M	LM	-	M	-
62	R	-	-	-	DKER	-	-
64	H	ANQHL	ANQH[L]	AQH	A	AQ	A
67	F	-	-	-	-	MYILVF	M**Y**ILVF
69	P	-	-	-	LP	-	-
73	E	-	-	-	DEA	-	-
76	Q	-	-	-	RDAQ	-	-
79	L	-	-	-	FVL	-	-
82	N	ANST	[A]NST	NT	-	NT	-
83	A	-	-	-	NSA	-	-
85	S	-	-	-	ANGS	-	-
86	F	MFWY	MFWY	**M**Y	M	M	M
88	C	-	-	-	RSC	-	-
93	H	-	-	-	YH	-	-
96	G	-	-	-	AG	-	-
98	K	-	-	-	RK	-	-
100	H	HMFWY	X **CW** (hydrophobic/aromatic)	WYC	C	**C**W	C
101	L	-	-	-	ISL	-	-
102	G	-	-	-	RG	-	-
104	R	AGR	AGR	AGR	G	-	G
105	P	ANDGPSTV	X **IVW** (hydrophobic/aromatic)	**W**IFV	F	W	F
108	K	-	-	-	**- (Q)**	**Q**K	**Q**K
109	Q	-	-	-	HNQ	ML**Q**H	ML**Q**H
110	Y	-	-	-	VFY	-	-
112	T	-	-	-	-	STAG	STAG
113	L	ARNDQEKMSTV	**A**RNDQEKMST**V**I**P**HG	PKA	A	A	A
114	E	-	-	-	RDE	-	-
116	Q	ARNQEIKMTV	X **SM**	MSNQ	S	MQ	S
117	L	-	-	-	IVL	-	ML
120	L	-	-	-	YML	-	-
125	F	-	-	-	LF	-	-
129	N	-	-	-	EDHQN	-	**N**H
130	A	-	-	-	GA	-	-
131	L	-	-	-	-	-	ML
134	L	ARILKMFWYV	X **RLVIW**	WMRV	M	M**V**	M
135	S	ARNQHKST	A**R**NQHKST	SKQ	Q	-	Q
137	V	-	-	-	LIAV		AV
138	G	AHKMFSYW	X **RCA**	C**R**A	C **(G)**	G**R**	C**G**
139	H	ARQHLKY	X **T**	NIVG	I	**V**N	I
140	F	-	-	-	YF	-	FY
145	V	-	-	-	AV	-	-
147	E	ARQEHLKMY	X **W** (bulky/hydrophobic)	SFWL	L	F	L
149	Q	-	-	-	RQ	-	-
151	H	AQHKIL	X **MRK**	RSML	L	SL	L
162	T	-	-	-	QT	-	-
164	D	-	-	DA	-	A	A
167	P	-	-	-	SP	-	-
170	L	-	-	-	FMIL **(V)**	**L**AV	LA**V**
173	A	-	-	-	-	**GA**V	GAV
174	I	ARQELKM	X **WL** (aromatic/hydrophobic/basic)	WSLV	L	L**W**	L
175	E	-	-	-	NDE	-	-
177	F	ARLKM	X **GL**	SLA**K**	K	K	K
181	G	-	-	-	SG	-	-
183	E	-	-	-	GDE	-	-
189	G	-	-	-	VG	-	-
190	L	-	-	-	VIL	-	-
193	I	-	-	-	VMLI	-	-
197	L	-	-	-	IFAL	-	-
203	C	-	-	CS	-	-	-

Figure 9


A
Control
2mg/l Chlorsulfuron
2mg/l Ethametsulfuron
8678.110.1.2
PHP37707
Single copy
8678.109.1.13
PHP37707
Single copy
8678.110.1.8
PHP37707
Induced 9/9/08-9/15/08

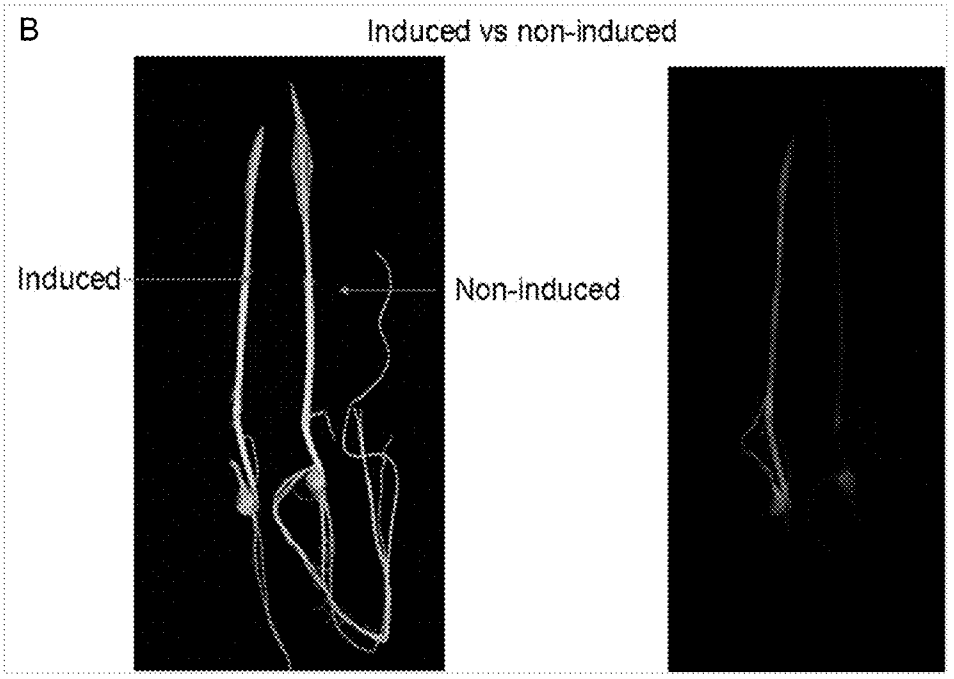

B
Induced vs non-induced
Induced
Non-induced

Figure 10

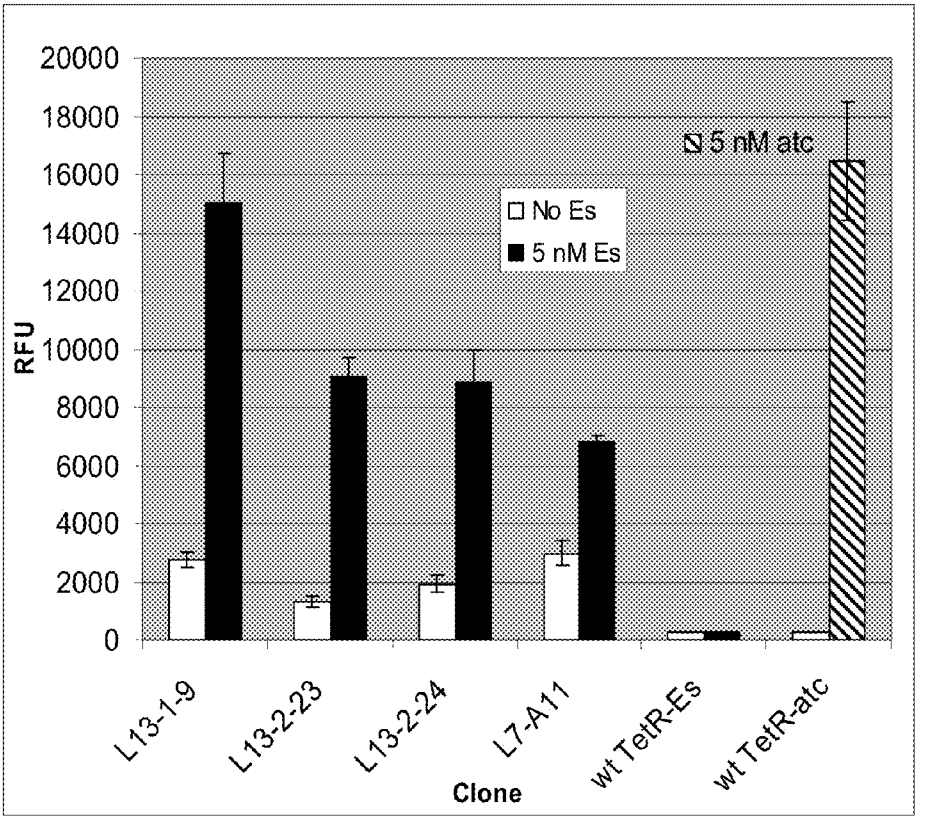

20000
18000
16000
14000
12000
10000
8000
6000
4000
2000
0
RFU
5 nM atc
No Es
5 nM Es
L13-1-9
L13-2-23
L13-2-24
L7-A11
wt TetR-Es
wt TetR-atc
Clone

Figure 11

SULFONYLUREA-RESPONSIVE REPRESSOR PROTEINS

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a Continuation of U.S. application Ser. No. 12/603,739 filed Oct. 22, 2009, now U.S. Pat. No. 8,257,956 which claims the benefit of U.S. Application Ser. No. 61/108,917 filed Oct. 28, 2008, which is herein incorporated by reference in its entirety.

FIELD OF THE INVENTION

The invention relates to the field of molecular biology, more particularly to the regulation of gene expression.

BACKGROUND

The tetracycline operon system, comprising repressor and operator elements, was originally isolated from bacteria. The operon system is tightly controlled by the presence of tetracycline, and self-regulates the level of expression of tetA and tetR genes. The product of tetA removes tetracycline from the cell. The product of tetR is the repressor protein that binds to the operator elements with a K_d of about 10 pM in the absence of tetracycline, thereby blocking expression or tetA and tetR.

This system has been modified to control expression of other polynucleotides of interest, and/or for use in other organisms, mainly for use in animal systems. Tet operon based systems have had limited use in plants, at least partially due to problems with the inducers which are typically antibiotic compounds, and sensitive to light.

There is a need to regulate expression of sequences of interest in organisms, compositions and methods to tightly regulate expression in response to sulfonylurea compounds are provided.

SUMMARY

Compositions and methods relating to the use of sulfonylurea-responsive repressors are provided. Compositions include polypeptides that specifically bind to an operator, wherein the specific binding is regulated by a sulfonylurea compound. Compositions also include polynucleotides encoding the polypeptides as well as constructs, vectors, prokaryotic and eukaryotic cells and eukaryotic organisms including plants and seeds comprising the polynucleotide and/or produced by the methods. Also provided are methods to provide a sulfonylurea-responsive repressor to a cell or organism and to regulate expression of a polynucleotide of interest in a cell or organism, including a plant or plant cell.

BRIEF DESCRIPTION OF DRAWINGS

FIG. **1**. Docking of tetracycline-Mg++ and the sulfonylurea compound Harmony® (thifensulfuron-methyl; Ts) into the binding pocket of class D TetR based on the crystal structure 1DU7 from the Protein Databank (PDB).

FIG. **2**. Vector map of an exemplary *E. coli*-based tetR expression vector, pVER7314. The replicon backbone is based on that of pBR322. The TetR ligand binding domain (LBD) is encoded flanked by SacI and AscI sites. KMsp172 and KMsp173 represent binding sites for the primers used for DNA sequencing of inserted tetR genes. rrnB T1 T2 is a strong transcriptional terminator to inhibit run around transcription and unregulated tetR expression.

FIG. **3**. Response of library 1 hits to 20 μg/ml thifensulfuron-methyl (Ts). *E. coli* KM3 cells harboring putative tetR hits L1-1 through L1-20 or wt tetR were replica plated onto M9 assay medium +/−20 μg/ml Ts, then incubated at 30° C. until blue/white colony discrimination was evident. At this time colonies were imaged and relative β-galactosidase activity determined based on degree of blue colony color.

FIG. **4**. Relative β-galactosidase activities of 45 putative library L4 hits against 0, 0.2 and 1.0 ppm ethametsulfuron (Es). Induced activity was measured using 5 μl of perforated whole cell mixture, and background activity was measured using 25 μl perforated cell mixture so that detectable activity could be measured in the same time frame for all treatments. Background activity values were multiplied by 10 in order to bring them into the display range of the graph. The right hand side of the graph contains the controls, wild type TetR and 1st round hit L1-9.

FIG. **5**. β-galactosidase induction in L7 hits with ethametsulfuron. Top hits from the L7 library were re-arrayed and tested in 96-well culture format for relative induction by 0.02 μg/ml and 0.2 μg/ml inducer (Es), and for background activity in the absence of inducer. Induced activity was assayed using 5 μl of perforated cell mixture, whereas 25 μl of cells was used to detect background activity. This allowed all detectable activities to be measured in the same time frame for all treatments. Background activity values were multiplied by ten to bring them into the range of the graph. The latter part of the graph shows the controls: 2nd round hits L4-89 and L4-120, and wt TetR(B) with ethametsulfuron; and wt TetR with 0.4 μM atc as cognate inducer for comparison (diagonally striped bar). Well ID's indicated with slanted text refer to that of the assay re-array whereas original clone ID's are indicated below in horizontal text.

FIG. **6**. Ethametsulfuron dose response of two EsR variants determined by transient expression in *Nicotiana benthamiana* leaves. Black bars represent wt TetR, grey bars represent EsR hit A11, and white bars represent EsR hit D01. The striped bar represents a no repressor control which indicates the maximum level of reporter expression in the assay.

FIG. **7**. DNA binding to tetOp in the absence or presence of ligand. Five pmol TetO or control DNA was mixed with the indicated amounts of repressor protein and inducer in complex buffer containing 20 mM Tris-HCl (pH8.0) and 10 mM EDTA.

FIG. **8**. Structures of exemplary registered sulfonylurea compounds.

FIG. **9**. Summary of source diversity, library design, and hit diversity and population bias for several generations of sulfonylurea repressor shuffling libraries. A dash ("-") indicates no amino acid diversity introduced at that position in that library. An X indicates that the library oligos were designed to introduce complete amino acid diversity (any of 20 amino acids) at that position in that library. Residues in bold indicate bias during selection with larger font size indicating a greater degree of bias in the selected population. Residues in parentheses indicate selected mutations. The phylogenetic diversity pool is derived from a broad family of 34 tetracycline repressor sequences.

FIG. **10**. Sulfonylurea depression of fluorescent reporter in maize callus (A) or plants (B).

FIG. **11**. β-galactosidase induction in exemplary L13 hits with ethametsulfuron.

DETAILED DESCRIPTION

Chemically regulated expression tools have proven valuable for studying gene function and regulation in many bio-

logical systems. These systems allow testing for the effect of expression of any gene of interest in a culture system or whole organism when the transgene cannot be specifically regulated, or continuously expressed due to negative consequences. These systems essentially provide the opportunity to do "pulse" or "pulse-chase" gene expression testing. A chemical switch-mediated expression system allows testing of genomic, proteomic, and/or metabolomic responses immediately following activation of the target gene. These types of tests cannot be done with constitutive, developmental, or tissue-specific expression systems. Chemical switch technologies may also provide a means for gene therapy.

Chemical switch systems can be commercially applied, such as in agricultural biotechnology. For agricultural purposes it is desired to be able to control the expression and/or genetic flow of transgenes in the environment, such as herbicide resistance genes, especially in cases where weedy relatives of the target crop exist. In addition, having a family of viable chemical switch mechanisms would enable trait inventory management from a single transgenic crop, for example, one production line could be used to deliver selected traits on customer demand via specific chemical activation. Additionally, hybrid seed production could be streamlined by using chemical control of hybrid maintenance.

The Tet repressor (TetR) based genetic switch system widely used in animal systems has had limited use in plant genetic systems, due in part to problems with the activator ligands. TetR has been redesigned to recognize commercially used sulfonylurea chemistry instead of tetracycline compounds, while retaining the ability to specifically bind tetracycline operator sequences. This was accomplished by modifying the Tet repressor ligand binding domain using rational protein modeling and DNA shuffling to recognize commercially used sulfonylurea compounds. Initial TetR shuffling and screening using a sensitive in vivo β-galactosidase assay led to specific recognition of the herbicide Harmony® (thifensulfuron-methyl) at 20 ppm in the growth medium, and loss of recognition of tetracycline. Upon testing with other sulfonylurea compounds, many of the hits reactive to Harmony® also responded to other SU compounds. In some cases, the hits had even better reactivity to related herbicides chlorsulfuron and ethametsulfuron (2 ppm). Further rounds of shuffling and screening of the TetR derivatives led to TetR variants that react robustly to 0.2 ppm chlorsulfuron and 0.02 ppm ethametsulfuron as measured using in vivo induction assays in *E. coli*. Top performing ethametsulfuron responsive SuR variants (EsRs) show induction capacity nearly equal to that of wild type class B TetR induction by anhydrotetracycline (atc) using similar inducer concentrations. These SuR molecules have no reactivity to tetracyclines, and wild type TetR(B) (SEQ ID NO: 2) has no reactivity to the sulfonylureas.

Compositions and methods relating to the use of sulfonylurea-responsive repressors are provided. Sulfonylurea-responsive repressors (SuRs) include any repressor polypeptide whose binding to an operator sequence is controlled by a ligand comprising a sulfonylurea compound. In some examples, the repressor binds specifically to the operator in the absence of a sulfonylurea ligand. In some examples, the repressor binds specifically to the operator in the presence of a sulfonylurea ligand. Repressors that bind to an operator in the presence of the ligand are sometimes called a reverse repressor. In some examples compositions include SuR polypeptides that specifically bind to a tetracycline operator, wherein the specific binding is regulated by a sulfonylurea compound. In some examples compositions include an isolated sulfonylurea repressor (SuR) polypeptide comprising at least one amino acid substitution to a wild type tetracycline repressor protein ligand binding domain wherein the SuR polypeptide, or a multimer thereof, specifically binds to a polynucleotide comprising an operator sequence, wherein repressor-operator binding is regulated by the absence or presence of a sulfonylurea compound. In some examples compositions included isolated sulfonylurea repressors comprising a ligand binding domain comprising at least one amino acid substitution to a wild type tetracycline repressor protein ligand binding domain fused to a heterologous operator DNA binding domain which specifically binds to a polynucleotide comprising the operator sequence or derivative thereof, wherein repressor-operator binding is regulated by the absence or presence of a sulfonylurea compound. Any operator DNA binding domain can be used, including but not limited to an operator DNA binding domain from repressors included tet, lac, trp, phd, arg, LexA, phiCh1 repressor, lambda C1 and Cro repressors, phage X repressor, MetJ, phir1t rro, phi434 C1 and Cro repressors, RafR, gal, ebg, uxuR, exuR, ROS, SinR, PurR, FruR, P22 C2, TetC, AcrR, Betl, Bm3R1, EnvR, QacR, MtrR, TcmR, Ttk, YbiH, YhgD, and mu Ner, or DNA binding domains in Interpro families including but not limited to IPR001647, IPR010982, and IPR011991.

In some examples compositions include an isolated sulfonylurea repressor (SuR) polypeptides comprising at least one amino acid substitution to a wild type tetracycline repressor protein wherein the SuR polypeptide, or a multimer thereof, specifically binds to a polynucleotide comprising a tetracycline operator sequence, wherein repressor-operator binding is regulated by the absence or presence of a sulfonylurea compound.

Wild type repressors include tetracycline class A, B, C, D, E, G, H, J and Z repressors. An example of the TetR(A) class is found on the Tn1721 transposon and deposited under GenBank accession X61307, crossreferenced under gi48198, with encoded protein accession CAA43639, crossreferenced under gi48195 and UniProt accession Q56321. An example of the TetR(B) class is found on the Tn10 transposon and deposited under GenBank accession X00694, crossreferenced under gi43052, with encoded protein accession CAA25291, crossreferenced under gi43052 and UniProt accession P04483. An example of the TetR(C) class is found on the pSC101 plasmid and deposited under GenBank Accession M36272, crossreferenced under gi150945, with encoded protein accession AAA25677, crossreferenced under gi150946. An example of the TetR(D) class is found in *Salmonella ordonez* and deposited under GenBank Accession X65876, crossreferenced under gi49073, with encoded protein accession CAA46707, crossreferenced under gi49075 and UniProt accessions P0ACT5 and P09164. An example of the TetR(E) class was isolated from *E. coli* transposon Tn10 and deposited under GenBank Accession M34933, crossreferenced under gi155019, with encoded protein accession AAA98409, crossreferenced under gi155020. An example of the TetR(G) class was isolated from *Vibrio anguillarium* and deposited under GenBank Accession S52438, crossreferenced under gi262928, with encoded protein accession AAB24797, crossreferenced under gi262929. An example of the TetR(H) class is found on plasmid pMV111 isolated from *Pasteurella multocida* and deposited under GenBank Accession U00792, crossreferenced under gi392871, with encoded protein accession AAC43249, crossreferenced under gi392872. An example of the TetR(J) class was isolated from *Proteus mirabilis* and deposited under GenBank Accession AF038993, crossreferenced under gi4104704, with encoded protein accession AAD12754, crossreferenced under

gi4104706. An example of the TetR(Z) class was found on plasmid pAGI isolated from *Corynebacterium glutamicum* and deposited under GenBank Accession AF121000, cross-referenced under gi4583389, with encoded protein accession AAD25064, crossreferenced under gi4583390. In some examples the wild type tetracycline repressor is a class B tetracycline repressor protein. In some examples the wild type tetracycline repressor is a class D tetracycline repressor protein.

In some examples the sulfonylurea repressor (SuR) polypeptides comprise an amino acid substitution in the ligand binding domain of a wild type tetracycline repressor protein. In class B and D wild type TetR proteins, amino acid residues 6-52 represent the DNA binding domain. The remainder of the protein is involved in ligand binding and subsequent allosteric modification. For class B TetR residues 53-207 represent the ligand binding domain, while residues 53-218 comprise the ligand binding domain for the class D TetR. In some examples the SuR polypeptides comprise an amino acid substitution in the ligand binding domain of a wild type TetR(B) protein. In some examples the SuR polypeptides comprise an amino acid substitution in the ligand binding domain of a wild type TetR(B) protein of SEQ ID NO: 1.

In some examples the isolated SuR polypeptides comprise an amino acid, or any combination of amino acids, corresponding to equivalent amino acid positions selected from the amino acid diversity shown in FIG. **9**, wherein the amino acid residue position shown in FIG. **9** corresponds to the amino acid numbering of a wild type TetR(B). In some examples the isolated SuR polypeptides comprise a ligand binding domain comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues shown in FIG. **9**, wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of a wild type TetR(B). In some examples the isolated SuR polypeptides comprise at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues shown in FIG. **9**, wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of a wild type TetR(B). In some examples the wild type TetR(B) is SEQ ID NO: 1.

In some examples the isolated SuR polypeptide comprises a ligand binding domain comprising an amino acid substitution at a residue position selected from the group consisting of position 55, 60, 64, 67, 82, 86, 100, 104, 105, 108, 113, 116, 134, 135, 138, 139, 147, 151, 170, 173, 174, 177 and any combination thereof, wherein the amino acid residue position and substitution corresponds to the equivalent position using the amino acid numbering of a wild type TetR(B). In some examples the isolated SuR polypeptide further comprises an amino acid substitution at a residue position selected from the group consisting of 109, 112, 117, 131, 137, 140, 164 and any combination thereof. In some examples the wild type TetR(B) is SEQ ID NO: 1.

In some examples the isolated SuR polypeptides comprise a ligand binding domain comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues selected from the group consisting of:

(a) M or L at amino acid residue position 55;
(b) A, L or M at amino acid residue position 60;
(c) A, N, Q, L or H at amino acid residue position 64;
(d) M, I, L, V, F or Y at amino acid residue position 67;
(e) N, S or T at amino acid residue position 82;
(f) F, M, W or Y at amino acid residue position 86;
(g) C, V, L, M, F, W or Y at amino acid residue position 100;
(h) R, A or G at amino acid residue position 104
(i) A, I, V, F or W at amino acid residue position 105;
(j) Q or K at amino acid residue position 108;
(k) A, M, H, K, T, P or V at amino acid residue position 113;
(l) I, L, M, V, R, S, N, P or Q at amino acid residue position 116;
(m) I, L, V, M, R, S or W at amino acid residue position 134;
(n) R, S, N, Q, K or A at amino acid residue position 135;
(o) A, C, G, H, I, V, R or T at amino acid residue position 138;
(p) A, G, I, V, M, W, N, R or T at amino acid residue position 139;
(q) I, L, V, F, W, T, S or R at amino acid residue position 147;
(r) M, L, W, Y, K, R or S at amino acid residue position 151;
(s) I, L, V or A at amino acid residue position 170;
(t) A, G or V at amino acid residue position 173;
(u) L, V, W, Y, H, R, K or S at amino acid residue position 174; and,
(v) A, G, I, L, Y, K, Q or S at amino acid residue position 177,

wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the isolated SuR polypeptides comprise at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues are selected from the amino acid residues listed in (a)-(v) above, wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the wild type TetR(B) is SEQ ID NO: 1.

In some examples the isolated SuR polypeptides selected for enhanced activity on chlorsulfuron comprise a ligand binding domain comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues are selected from the group consisting of:

(a) M at amino acid residue position 60;
(b) A or Q at amino acid residue position 64;
(c) M, F, Y, I, V or L at amino acid residue position 67;
(d) N or T at amino acid residue position 82;
(e) M at amino acid residue position 86;
(f) C or W at amino acid residue position 100;
(g) W at amino acid residue position 105;
(h) Q or K at amino acid residue position 108;
(i) M, Q, L or H at amino acid residue position 109;
(j) G, A, S or T at amino acid residue position 112;
(k) A at amino acid residue position 113;
(l) M or Q at amino acid residue position 116;
(m) M or V at amino acid residue position 134;
(n) G or R at amino acid residue position 138;
(o) N or V at amino acid residue position 139;

(p) F at amino acid residue position 147;
(q) S or L at amino acid residue position 151;
(r) A at amino acid residue position 164;
(s) A, L or V at amino acid residue position 170;
(t) A, G or V at amino acid residue position 173;
(u) L or W at amino acid residue position 174; and;
(v) K at amino acid residue position 177,

wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the isolated SuR polypeptides comprise at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues are selected from the amino acid residues listed in (a)-(v) above, wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the wild type TetR(B) is SEQ ID NO: 1.

In some examples the isolated SuR polypeptides selected for enhanced activity on ethametsulfuron comprise a ligand binding domain comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% the amino acid residues are selected from the group consisting of:

(a) M or L at amino acid residue position 55;
(b) A at amino acid residue position 64;
(c) M, Y, F, I, L or V at amino acid residue position 67;
(d) M at amino acid residue position 86;
(e) C at amino acid residue position 100;
(f) G at amino acid residue position 104;
(g) F at amino acid residue position 105;
(h) Q or K at amino acid residue position 108;
(i) Q, M, L or H at amino acid residue position 109;
(j) S, T, G or A at amino acid residue position 112;
(k) A at amino acid residue position 113;
(l) S at amino acid residue position 116;
(m) M or L at amino acid residue position 117;
(n) M or L at amino acid residue position 131;
(o) M at amino acid residue position 134;
(p) Q at amino acid residue position 135;
(q) A or V at amino acid residue position 137;
(r) C or G at amino acid residue position 138;
(s) I at amino acid residue position 139;
(t) F or Y at amino acid residue position 140;
(u) L at amino acid residue position 147;
(v) L at amino acid residue position 151;
(w) A at amino acid residue position 164;
(x) V, A or L at amino acid residue position 170;
(y) G, A or V at amino acid residue position 173
(z) L at amino acid residue position 174; and,
(aa) N or K at amino acid residue position 177,

wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the isolated SuR polypeptides comprise at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% the amino acid residues are selected from the amino acid residues listed in (a)-(aa) above, wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the wild type TetR(B) is SEQ ID NO: 1.

In some examples the isolated SuR polypeptide has at least about 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to the ligand binding domain of a wild type TetR(B) exemplified by amino acid residues 53-207 of SEQ ID NO: 1, wherein the sequence identity is determined over the full length of the ligand binding domain using a global alignment method. In some examples the global alignment method uses the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix.

In some examples the isolated SuR polypeptide has at least about 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to a wild type TetR(B) exemplified by SEQ ID NO: 1, wherein the sequence identity is determined over the full length of the polypeptide using a global alignment method. In some examples the global alignment method uses the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix.

Compositions include isolated SuR polypeptides having at least about 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to the ligand binding domain of an SuR polypeptide selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243, wherein the sequence identity is determined over the full length of the ligand binding domain using a global alignment method. In some examples the global alignment method uses the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix.

In some examples the isolated SuR polypeptide have at least about 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to an SuR polypeptide selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243, wherein the sequence identity is determined over the full length of the polypeptide using a global alignment method. In some examples the global alignment method uses the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO:220) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO:220) to generate a BLAST bit score of at least 374, wherein the BLAST alignment used the BLOSUM62 matrix,

a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO:220) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO:220) to generate a percent sequence identity of at least 88% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO:220) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO:220) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a BLAST bit score of at least 387, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a percent sequence identity of at least 92% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a BLAST bit score of at least 393, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030,

1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a BLAST similarity score of at least 1006, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-02 (SEQ ID NO: 3) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-02 (SEQ ID NO: 3) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-02 (SEQ ID NO: 3) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-02 (SEQ ID NO: 3) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-02 (SEQ ID NO: 3) to generate a BLAST e-value score of at least e-112, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST bit score of at least 388, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. n some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a percent sequence identity of at least 93% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST similarity score of at least 996, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST e-value score of at least e-111, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap

extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a percent sequence identity of at least 93% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a BLAST e-value score of at least e-111, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-44 (SEQ ID NO: 13) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-44 (SEQ ID NO: 13) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-44 (SEQ ID NO: 13) to generate a percent sequence identity of at least 93% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-44 (SEQ ID NO: 13) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-44 (SEQ ID NO: 13) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST bit score of at least 381, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using

the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST similarity score of at least 978 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST e-value score of at least e-108, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3H02 (SEQ ID NO: 94) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3H02 (SEQ ID NO: 94) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3H02 (SEQ ID NO: 94) to generate a percent sequence identity of at least 90% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3H02 (SEQ ID NO: 94) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3H02 (SEQ ID NO: 94) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST bit score of at least 368, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST similarity score of at least 945 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113,

e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST e-value score of at least e-105, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L10-84(B12) (SEQ ID NO: 1230) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L10-84 (B12) (SEQ ID NO: 1230) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L10-84(B12) (SEQ ID NO: 1230) to generate a percent sequence identity of at least 86% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L10-84 (B12) (SEQ ID NO: 1230) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L10-84 (B12) (SEQ ID NO: 1230) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST bit score of at least 320, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a percent sequence identity of at least 86% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST similarity score of at least 819 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the SuR polypeptides comprise an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST e-value score of at least e-90, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243.

In some examples the isolated SuR polypeptides comprise a ligand binding domain from a polypeptide selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated SuR polypeptides comprise an amino acid sequence selected from

the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated SuR polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243, and the sulfonylurea compound is selected from the group consisting of a chlorsulfuron, an ethametsulfuron, a metsulfuron, a sulfometuron, a tribenuron, a chlorimuron, a nicosulfuron, a rimsulfuron and a thifensulfuron.

In some examples the isolated SuR polypeptides have an equilibrium binding constant for a sulfonylurea compound greater than 0.1 nM and less than 10 μM. In some examples the isolated SuR polypeptide has an equilibrium binding constant for a sulfonylurea compound of at least 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM, 1 μM, 5 μM, 7 μM but less than 10 μM. In some examples the isolated SuR polypeptide has an equilibrium binding constant for a sulfonylurea compound of at least 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM but less than 1 μM. In some examples the isolated SuR polypeptide has an equilibrium binding constant for a sulfonylurea compound greater than 0 nM, but less than 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM, 1 μM, 5 μM, 7 μM or 10 μM. In some examples the sulfonylurea compound is a chlorsulfuron, an ethametsulfuron, a metsulfuron, a sulfometuron, a tribenuron, a chlorimuron, a nicosulfuron, a rimsulfuron and/or a thifensulfuron.

In some examples the isolated SuR polypeptides have an equilibrium binding constant for an operator sequence greater than 0.1 nM and less than 10 μM. In some examples the isolated SuR polypeptide has an equilibrium binding constant for an operator sequence of at least 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM, 1 μM, 5 μM, 7 μM but less than 10 μM. In some examples the isolated SuR polypeptide has an equilibrium binding constant for an operator sequence of at least 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM but less than 1 μM. In some examples the isolated SuR polypeptide has an equilibrium binding constant for an operator sequence greater than 0 nM, but less than 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM, 1 μM, 5 μM, 7 μM or 10 μM. In some examples the operator sequence is a Tet operator sequence. In some examples the Tet operator sequence is a TetR(A) operator sequence, a TetR(B) operator sequence, a TetR(D) operator sequence, TetR(E) operator sequence, a TetR(H) operator sequence, or a functional derivative thereof.

The isolated SuR polypeptides specifically bind to a sulfonylurea compound. Sulfonylurea molecules comprise a sulfonylurea moiety (—S(O)2NHC(O)NH(R)—). In sulfonylurea herbicides the sulfonyl end of the sulfonylurea moiety is connected either directly or by way of an oxygen atom or an optionally substituted amino or methylene group to a typically substituted cyclic or acyclic group. At the opposite end of the sulfonylurea bridge, the amino group, which may have a substituent such as methyl (R being CH3) instead of hydrogen, is connected to a heterocyclic group, typically a symmetric pyrimidine or triazine ring, having one or two substituents such as methyl, ethyl, trifluoromethyl, methoxy, ethoxy, methylamino, dimethylamino, ethylamino and the halogens. Sulfonylurea herbicides can be in the form of the free acid or a salt. In the free acid form the sulfonamide nitrogen on the bridge is not deprotonated (i.e., —S(O)2NHC(O)NH(R)—), while in the salt form the sulfonamide nitrogen atom on the bridge is deprotonated (i.e., —S(O)2NC(O)NH(R)—), and a cation is present, typically of an alkali metal or alkaline earth metal, most commonly sodium or potassium. Sulfonylurea compounds include, for example, compound classes such as pyrimidinylsulfonylurea compounds, triazinylsulfonylurea compounds, thiadiazolylurea compounds, and pharmaceuticals such as antidiabetic drugs, as well as salts and other derivatives thereof. Examples of pyrimidinylsulfonylurea compounds include amidosulfuron, azimsulfuron, bensulfuron, bensulfuron-methyl, chlorimuron, chlorimuron-ethyl, cyclosulfamuron, ethoxysulfuron, flazasulfuron, flucetosulfuron, flupyrsulfuron, flupyrsulfuron-methyl, foramsulfuron, halosulfuron, halosulfuron-methyl, imazosulfuron, mesosulfuron, mesosulfuron-methyl, nicosulfuron, orthosulfamuron, oxasulfuron, primisulfuron, prim isulfuron-methyl, pyrazosulfuron, pyrazosulfuron-ethyl, rimsulfuron, sulfometuron, sulfometuron-methyl, sulfosulfuron, trifloxysulfuron and salts and derivatives thereof. Examples of triazinylsulfonylurea compounds include chlorsulfuron, cinosulfuron, ethametsulfuron, ethametsulfuron-methyl, iodosulfuron, iodosulfuron-methyl, metsulfuron, metsulfuron-methyl, prosulfuron, thifensulfuron, thifensulfuron-methyl, triasulfuron, tribenuron, tribenuron-methyl, triflusulfuron, triflusulfuron-methyl, tritosulfuron and salts and derivatives thereof. Examples of thiadiazolylurea compounds include buthiuron, ethidimuron, tebuthiuron, thiazafluoron, thidiazuron and salts and derivatives thereof. Examples of antidiabetic drugs include acetohexamide, chlorpropamide, tolbutamide, tolazamide, glipizide, gliclazide, glibenclamide (glyburide), gliquidone, glimepiride and salts and derivatives thereof. In some examples the isolated SuR polypeptides specifically bind to more than one sulfonylurea compound. In some examples the sulfonylurea compound is selected from the group consisting of chlorsulfuron, ethametsulfuron-methyl, metsulfuron-methyl, thifensulfuron-methyl, sulfometuron-methyl, tribenuron-methyl, chlorimuron-ethyl, nicosulfuron, and rimsulfuron.

Compositions also include isolated polynucleotides encoding SuR polypeptides that specifically bind to a tetracycline operator, wherein the specific binding is regulated by a sulfonylurea compound. In some examples the isolated polynucleotides encode sulfonylurea repressor (SuR) polypeptides comprising an amino acid substitution in the ligand binding domain of a wild type tetracycline repressor protein. In class B and D wild type TetR proteins, amino acid residues 6-52 represent the DNA binding domain. The remainder of the protein is involved in ligand binding and subsequent allosteric modification. For class B TetR residues 53-207 represent the ligand binding domain, while residues 53-218 comprise the ligand binding domain for the class D TetR. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid substitution in the ligand binding domain of a wild type TetR(B) protein. In some examples the polynucleotides encode SuR polypeptides comprising an amino acid substitution in the ligand binding domain of a wild type TetR(B) protein of SEQ ID NO: 1.

In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid, or any combination of amino acids, selected from the amino acid diversity shown in FIG. **9**, wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of a wild type TetR(B) exemplified by SEQ ID NO: 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising a ligand binding domain comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues shown in FIG. **9**, wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the iso-

lated polynucleotides encode SuR polypeptides comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues shown in FIG. **9**, wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the wild type TetR(B) is SEQ ID NO: 1.

In some examples the isolated polynucleotides encode SuR polypeptides comprising a ligand binding domain comprising an amino acid substitution at a residue position selected from the group consisting of position 55, 60, 64, 67, 82, 86, 100, 104, 105, 108, 113, 116, 134, 135, 138, 139, 147, 151, 170, 173, 174, 177 and any combination thereof, wherein the amino acid residue position and substitution corresponds to the equivalent position using the amino acid numbering of a wild type TetR(B). In some examples the isolated polynucleotides encode SuR polypeptides further comprising an amino acid substitution at a residue position selected from the group consisting of 109, 112, 117, 131, 137, 140, 164 and any combination thereof. In some examples the wild type TetR(B) polypeptide sequence is SEQ ID NO: 1.

In some examples the isolated polynucleotides encode SuR polypeptides having a ligand binding domain comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues are selected from the group consisting of:

(a) M or L at amino acid residue position 55;
(b) A, L or M at amino acid residue position 60;
(c) A, N, Q, L or H at amino acid residue position 64;
(d) M, I, L, V, F or Y at amino acid residue position 67;
(e) N, S or T at amino acid residue position 82;
(f) F, M, W or Y at amino acid residue position 86;
(g) C, V, L, M, F, W or Y at amino acid residue position 100;
(h) R, A or G at amino acid residue position 104
(i) A, I, V, F or W at amino acid residue position 105;
(j) Q or K at amino acid residue position 108;
(k) A, M, H, K, T, P or V at amino acid residue position 113;
(l) I, L, M, V, R, S, N, P or Q at amino acid residue position 116;
(m) I, L, V, M, R, S or W at amino acid residue position 134;
(n) R, S, N, Q, K or A at amino acid residue position 135;
(o) A, C, G, H, I, V, R or T at amino acid residue position 138;
(p) A, G, I, V, M, W, N, R or T at amino acid residue position 139;
(q) I, L, V, F, W, T, S or R at amino acid residue position 147;
(r) M, L, W, Y, K, R or S at amino acid residue position 151;
(s) I, L, V or A at amino acid residue position 170;
(t) A, G or V at amino acid residue position 173;
(u) L, V, W, Y, H, R, K or S at amino acid residue position 174; and,
(v) A, G, I, L, Y, K, Q or S at amino acid residue position 177,

wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the isolated SuR polynucleotides encode SuR polypeptides comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues are selected from the amino acid residues listed in (a)-(v) above, wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the wild type TetR(B) is SEQ ID NO: 1.

In some examples the isolated polynucleotides encode SuR polypeptides selected for enhanced activity on chlorsulfuron having a ligand binding domain comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues are selected from the group consisting of:

(a) M at amino acid residue position 60;
(b) A or Q at amino acid residue position 64;
(c) M, F, Y, I, V or L at amino acid residue position 67;
(d) N or T at amino acid residue position 82;
(e) M at amino acid residue position 86;
(f) C or W at amino acid residue position 100;
(g) W at amino acid residue position 105;
(h) Q or K at amino acid residue position 108;
(i) M, Q, L or H at amino acid residue position 109;
(j) G, A, S or T at amino acid residue position 112;
(k) A at amino acid residue position 113;
(l) M or Q at amino acid residue position 116;
(m) M or V at amino acid residue position 134;
(n) G or R at amino acid residue position 138;
(o) N or V at amino acid residue position 139;
(p) F at amino acid residue position 147;
(q) S or L at amino acid residue position 151;
(r) A at amino acid residue position 164;
(s) A, L or V at amino acid residue position 170;
(t) A, G or V at amino acid residue position 173;
(u) L or W at amino acid residue position 174; and;
(v) K at amino acid residue position 177,

wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the isolated SuR polynucleotides encode SuR polypeptides comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues are selected from the amino acid residues listed in (a)-(v) above wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the wild type TetR(B) is SEQ ID NO: 1.

In some examples the isolated polynucleotides encode SuR polypeptides selected for enhanced activity on ethametsulfuron having a ligand binding domain comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues are selected from the group consisting of:

(a) M or L at amino acid residue position 55;
(b) A at amino acid residue position 64;
(c) M, Y, F, I, L or V at amino acid residue position 67;
(d) M at amino acid residue position 86;
(e) C at amino acid residue position 100;
(f) G at amino acid residue position 104;
(g) F at amino acid residue position 105;
(h) Q or K at amino acid residue position 108;
(i) Q, M, L or H at amino acid residue position 109;
(j) S, T, G or A at amino acid residue position 112;
(k) A at amino acid residue position 113;

(l) S at amino acid residue position 116;
(m) M or L at amino acid residue position 117;
(n) M or L at amino acid residue position 131;
(o) M at amino acid residue position 134;
(p) Q at amino acid residue position 135;
(q) A or V at amino acid residue position 137;
(r) C or G amino acid residue 138;
(s) I at amino acid residue position 139;
(t) F or Y at amino acid residue position 140;
(u) L at amino acid residue position 147;
(v) L at amino acid residue position 151;
(w) A at amino acid residue position 164;
(x) V, A or L at amino acid residue position 170;
(y) G, A or V at amino acid residue position 173;
(z) L at amino acid residue position 174; and,
(aa) N or K at amino acid residue position 177,

wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR(B). In some examples the isolated SuR polynucleotides encode SuR polypeptides comprising at least 10%, 20%, 30%, 40%, 50%, 55%, 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% or 100% of the amino acid residues are selected from the amino acid residues listed in (a)-(aa) above, wherein the amino acid residue position corresponds to the equivalent position using the amino acid numbering of wild type TetR (B). In some examples the wild type TetR(B) is SEQ ID NO: 1.

In some examples the isolated polynucleotides encode SuR polypeptides having at least about 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to the ligand binding domain shown as amino acid residues 53-207 of SEQ ID NO: 1, wherein the sequence identity is determined over the full length of the ligand binding domain using a global alignment method. In some examples the global alignment method is GAP, wherein the default parameters are for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix.

In some examples the isolated polynucleotides encode SuR polypeptides having at least about 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to SEQ ID NO: 1, wherein the sequence identity is determined over the full length of the polypeptide using a global alignment method. In some examples the global alignment method is GAP, wherein the default parameters are for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix.

In some examples the isolated polynucleotides include nucleic acid sequences that selectively hybridize under stringent hybridization conditions to a polynucleotide encoding a SuR polypeptide. Polynucleotides that selectively hybridize are polynucleotides which bind to a target sequence at a level of at least 2-fold over background as compared to hybridization to a non-target sequence. Stringent conditions are sequence-dependent and condition-dependent. Typical stringent conditions are those in which the salt concentration about 0.01 to 1.0 M at pH 7.0-8.3 at 30° C. for short probes (e.g., 10 to 50 nucleotides) or about 60° C. for long probes (e.g., greater than 50 nucleotides). Stringent conditions may include formamide or other destabilizing agents. Exemplary moderate stringency conditions include hybridization in 40 to 45% formamide, 1 M NaCl, 1% SDS at 37° C., and a wash in 0.5× to 1×SSC at 55 to 60° C. Exemplary high stringency conditions include hybridization in 50% formamide, 1 M NaCl, 1% SDS at 37° C., and a wash in 0.1×SSC at 60 to 65° C.

Specificity is impacted by post-hybridization wash conditions, typically via ionic strength and temperature. For DNA-DNA hybrids, the T_m can be approximated from the equation of Meinkoth and Wahl, (1984) *Anal. Biochem.* 138:267-284: T_m=81.5° C.+16.6 (log M)+0.41 (% GC)−0.61 (% form)−500/L; where M is the molarity of monovalent cations, % GC is the percentage of guanosine and cytosine nucleotides in the DNA, % form is the percentage of formamide in the hybridization solution, and L is the length of the hybrid in base pairs. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology—Hybridization with Nucleic Acid Probes*, Part I, Chapter 2 "Overview of principles of hybridization and the strategy of nucleic acid probe assays", Elsevier, New York (1993); and *Current Protocols in Molecular Biology*, Chapter 2, Ausubel, et al., Eds., Greene Publishing and Wiley-Interscience, New York (1995). In some examples, the isolated polynucleotides encoding SuR polypeptides specifically hybridize to a polynucleotide of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247 under moderately stringent conditions or under highly stringent conditions.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO: 220) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO: 220) to generate a BLAST bit score of at least 374, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO:220) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO:220) to generate a percent sequence identity of at least 88% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypep-

tides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO:220) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 600, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200, wherein BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-1A04 (SEQ ID NO:220) to generate a BLAST e-value score of at least e-60, e-70, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a BLAST bit score of at least 387, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a percent sequence identity of at least 92% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 600, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200, wherein BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-22 (SEQ ID NO: 7) to generate a BLAST e-value score of at least e-60, e-70, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a BLAST bit score of at least 393, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a BLAST similarity score of at least 1006, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence pen-

alty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-29 (SEQ ID NO: 10) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-02 (SEQ ID NO: 3) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-02 (SEQ ID NO: 3) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-02 (SEQ ID NO: 3) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-02 (SEQ ID NO: 3) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-02 (SEQ ID NO: 3) to generate a BLAST e-value score of at least e-112, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST bit score of at least 388, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a percent sequence identity of at least 93% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST similarity score of at least 996 wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap

extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-07 (SEQ ID NO: 4) to generate a BLAST e-value score of at least e-111, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a percent sequence identity of at least 93% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-20 (SEQ ID NO: 6) to generate a BLAST e-value score of at least e-111, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-44 (SEQ ID NO: 13) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-44 (SEQ ID NO: 13) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-44 (SEQ ID NO: 13) to generate a percent sequence identity of at least 93% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-44 (SEQ ID NO: 13) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L1-44 (SEQ ID NO: 13) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST bit score of at least 381, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST similarity score of at least 978, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3A09 (SEQ ID NO: 1228) to generate a BLAST e-value score of at least e-108, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3H02 (SEQ ID NO: 94) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3H02 (SEQ ID NO: 94) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3H02 (SEQ ID NO: 94) to generate a percent sequence identity of at least 90% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3H02 (SEQ ID NO: 94) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L6-3H02 (SEQ ID NO: 94) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some

examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST bit score of at least 368, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST similarity score of at least 945, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L7-4E03 (SEQ ID NO: 1229) to generate a BLAST e-value score of at least e-105, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L10-84 (B12) (SEQ ID NO: 1230) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L10-84 (B12) (SEQ ID NO: 1230) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L10-84 (B12) (SEQ ID NO: 1230) to generate a percent sequence identity of at least 86% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L10-84(B12) (SEQ ID NO: 1230) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L10-84 (B12) (SEQ ID NO: 1230) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST bit score of at least 200, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, or 750, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode an SuR polypeptide comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST bit score of at least 320, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence

that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a percent sequence identity of at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a percent sequence identity of at least 86% sequence identity, wherein the sequence identity is determined by BLAST alignment using the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the percent sequence identity is determined using a global alignment method using the GAP algorithm with default parameters for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST similarity score of at least 400, 425, 450, 475, 500, 525, 550, 575, 600, 625, 650, 675, 700, 750, 800, 850, 900, 910, 920, 930, 940, 950, 960, 970, 980, 990, 1000, 1010, 1020, 1030, 1040, 1050, 1060, 1070, 1080, 1090, 1100, 1110, 1120, 1130, 1140, 1150, 1160, 1170, 1180, 1190, or 1200, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST similarity score of at least 819, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST e-value score of at least e-60, e-70, e-75, e-80, e-85, e-90, e-95, e-100, e-105, e-106, e-107, e-108, e-109, e-110, e-111, e-112, e-113, e-114, e-115, e-116, e-117, e-118, e-119, e-120, or e-125, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence that can be optimally aligned with a polypeptide sequence of L13-46 (SEQ ID NO: 1231) to generate a BLAST e-value score of at least e-90, wherein the BLAST alignment used the BLOSUM62 matrix, a gap existence penalty of 11, and a gap extension penalty of 1. In some examples isolated polynucleotide encodes a polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated polynucleotides encode an SuR polypeptide comprising a ligand binding domain from a polypeptide selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the isolated polynucleotides encode SuR polypeptides comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the encoded SuR polypeptide is selected from the group consisting of SEQ ID NO: 3-401, 1206-1213, 1228-1233, or 1240-1243, and the sulfonylurea compound is selected from the group consisting of chlorsulfuron, ethametsulfuron-methyl, metsulfuron-methyl, sulfometuron-methyl, and thifensulfuron-methyl. In some examples the isolated polynucleotides comprise a polynucleotide sequence of SEQ ID NO: 434-832, 1214-1221, 1234-1239, or 1244-1247, or the complementary polynucleotide thereof.

In some examples the isolated SuR polynucleotide encodes an SuR polypeptide having an equilibrium binding constant for a sulfonylurea compound greater than 0.1 nM and less than 10 μM. In some examples the encoded SuR polypeptide has an equilibrium binding constant for a sulfonylurea compound of at least 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM, 1 μM, 5 μM, 7 μM but less than 10 μM. In some examples the encoded SuR polypeptide has an equilibrium binding constant for a sulfonylurea compound of at least 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM but less than 1 μM. In some examples the encoded SuR polypeptide has an equilibrium binding constant for a sulfonylurea compound greater than 0 nM, but less than 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM, 1 μM, 5 μM, 7 μM, or 10 μM. In some examples the sulfonylurea compound is a chlorsulfuron, an ethametsulfuron, a metsulfuron, a sulfometuron, and/or a thifensulfuron compound.

In some examples the isolated SuR polynucleotide encodes an SuR polypeptide having an equilibrium binding constant for an operator sequence greater than 0.1 nM and less than 10 μM. In some examples the encoded SuR polypeptide has an equilibrium binding constant for an operator sequence of at least 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM, 1 μM, 5 μM, 7 μM but less than 10 μM. In some examples the encoded SuR polypeptide has an equilibrium binding constant for an operator sequence of at least 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM but less than 1 μM. In some examples the encoded SuR polypeptide has an equilibrium binding constant for an operator sequence greater than 0 nM, but less than 0.1 nM, 0.5 nM, 1 nM, 10 nM, 50 nM, 100 nM, 250 nM, 500 nM, 750 nM, 1 μM, 5 μM, 7 μM or 10 μM. In some examples the operator sequence is a Tet operator sequence. In some examples the Tet operator sequence is a TetR(A) operator sequence, a TetR(B) operator sequence, a TetR(D) operator sequence, TetR(E) operator sequence, a TetR(H) operator sequence or a functional derivative thereof.

In some examples the isolated polynucleotides encoding SuR polypeptides comprise codon composition profiles representative of codon preferences for particular host cells, or host cell organelles. In some examples the isolated polynucleotides comprise prokaryote preferred codons. In some examples the isolated polynucleotides comprise bacteria preferred codons. In some examples the bacteria is *E. coli* or *Agrobacterium*. In some examples the isolated polynucleotides comprise plastid preferred codons. In some examples the isolated polynucleotides comprise eukaryote preferred codons. In some examples the isolated polynucleotides comprise nuclear preferred codons. In some examples the isolated polynucleotides comprise plant preferred codons. In some examples the isolated polynucleotides comprise monocotyledonous plant preferred codons. In some examples the isolated polynucleotides comprise corn, rice, sorghum, barley, wheat, rye, switch grass, turf grass and/or oat preferred codons. In some examples the isolated polynucleotides comprise dicotyledonous plant preferred codons. In some

examples the isolated polynucleotides comprise soybean, sunflower, safflower, *Brassica*, alfalfa, *Arabidopsis*, tobacco and/or cotton preferred codons. In some examples the isolated polynucleotides comprise yeast preferred codons. In some examples the isolated polynucleotides comprise mammalian preferred codons. In some examples the isolated polynucleotides comprise insect preferred codons.

Compositions also include isolated polynucleotides fully complementary to a polynucleotide encoding an SuR polypeptide, expression cassettes, replicons, vectors, T-DNAs, DNA libraries, host cells, tissues and/or organisms comprising the polynucleotides encoding the SuR polypeptides and/or complements or derivatives thereof. In some examples a DNA library comprising a population of polynucleotides which encode a population of SuR polypeptide variants is provided. In some examples the polynucleotide is stably incorporated into a genome of the host cell, tissue and/or organism. In some examples the host cell is a prokaryote, including *E. coli* and *Agrobacterium* strains. In some examples the host is a eukaryote, including for example yeast, insects, plants and mammals.

Methods using the compositions are further provided. In one example methods of regulating transcription of a polynucleotide of interest in a host cell are provided, the methods comprising: providing a cell comprising the polynucleotide of interest operably linked to a promoter comprising at least one tetracycline operator sequence; providing an SuR polypeptide and, providing a sulfonylurea compound, thereby regulating transcription of the polynucleotide of interest. Any host cell can be used, including for example prokaryotic cells such as bacteria, and eukaryotic cells, including yeast, plant, insect, and mammalian cells. In some examples providing the SuR polypeptide comprises contacting the cell with an expression cassette comprising a promoter functional in the cell operably linked to a polynucleotide that encodes the SuR polypeptide.

Methods for generating and selecting diversified libraries to produce additional SuR polynucleotides, including polynucleotides encoding SuR polypeptides with improved and/or enhanced characteristics, e.g., altered binding constants for sulfonylurea compounds and/or the target DNA operator sequence and/or increased stability, all based upon selection of a polynucleotide constituent of the library for the new or improved activities are also provided. In some examples at least one library or population of oligonucleotides designed to introduce sequence modifications and/or diversity to a wild type or modified TetR ligand binding domain polypeptide is provided. In some examples the library or population is designed to introduce modifications and/or diversity to a wild type or modified TetR polypeptide. In some examples, the library or population introduce at least one modification as exemplified in FIG. **9**. In some examples the library or population comprises at least 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, or greater distinct oligonucleotides. In some examples the library or population comprises oligonucleotides selected from the oligonucleotides shown in at least one of Table 2, 9, 12, 13, 15, 17, 19 or a combination thereof. In some examples the library or population comprises one or more oligonucleotides selected from the group consisting of SEQ ID NO: 833-882, 885-986, 987-059, 1060-1083, 1084-1124, 1125-1154, 1159-1205.

In some examples the sulfonylurea compound is an ethametsulfuron. In some examples the ethametsulfuron is provided at a concentration of about 0.001, 0.002, 0.003, 0.004, 0.005, 0.006, 0.007, 0.008, 0.009, 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 200 or 500 μg/ml. In some examples the SuR polypeptide has a ligand binding domain having at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to an SuR polypeptide of SEQ ID NO: 205-401, 1206-1213, or 1228-1233, wherein the sequence identity is determined over the full length of the polypeptide using a global alignment method. In some examples the global alignment method is GAP, wherein the default parameters are for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix. In some examples the polypeptide has a ligand binding domain from a SuR polypeptide selected from the group consisting of SEQ ID NO: 205-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 205-401, 1206-1213, 1228-1233, or 1240-1243. In some examples the polypeptide is encoded by a polynucleotide of SEQ ID NO: 636-832, 1214-1221, 1234-1239, or 1244-1247.

In some examples the sulfonylurea compound is chlorsulfuron. In some examples the chlorsulfuron is provided at a concentration of about 0.01, 0.02, 0.03, 0.04, 0.05, 0.06, 0.07, 0.08, 0.09, 0.10, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95, 1.0, 1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 5.5, 6.0, 6.5, 7.0, 7.5, 8.0, 8.5, 9.0, 9.5, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 100, 200 or 500 μg/ml. In some examples the SuR polypeptide has a ligand binding domain having at least 50% 60%, 65%, 66%, 67%, 68%, 69%, 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98% or 99% sequence identity to an SuR polypeptide of SEQ ID NO: 14-204, wherein the sequence identity is determined over the full length of the polypeptide using a global alignment method. In some examples the global alignment method is GAP, wherein the default parameters are for an amino acid sequence % identity and % similarity using GAP Weight of 8 and Length Weight of 2 and the BLOSUM62 scoring matrix. In some examples the polypeptide has a ligand binding domain from a SuR polypeptide selected from the group consisting of SEQ ID NO: 14-204. In some examples the polypeptide is selected from the group consisting of SEQ ID NO: 14-204. In some examples the polypeptide is encoded by a polynucleotide of SEQ ID NO: 445-635.

The ability to capture value in various seed markets will require development of technology for controlling engineered trait distribution. One option is a trait deactivation/activation system using a chemically-regulated gene switch. To date no such system exists, in large part because of the lack of relevant chemistry, for example agricultural-compatible and/or pharmaceutical-based chemistry, that can be used as a ligand for a sensitive gene switch technology.

To develop an agricultural chemical-based ligand gene switch, TetR was modified using protein modeling, DNA shuffling, and a highly sensitive screening mechanism to produce a repressor that specifically recognizes sulfonylurea compounds. For agricultural applications, sulfonylurea compounds are phloem mobile and commercially available, thereby providing a good basis for use as switch ligand chemistry. Following three rounds of modeling and DNA shuffling, repressors that recognize SU chemistry nearly as well as wild

type TetR recognizes cognate inducers and yet are totally specific to sulfonylurea chemistry have been generated. These polypeptides comprise true sulfonylurea repressors (SuRs), which have been validated in planta using a newly developed transient assay system to demonstrate functionality of the SuR switch system. While exemplified in an agricultural context, these methods and compositions can be used in a wide variety of other settings and organisms.

In general, a chemical switch system wherein the chemical used penetrates rapidly and is perceived by all cell types in the organism, but does not perturb any endogenous regulatory networks will be most useful. Other important aspects have to do with the behavior of the sensor component, for example the stringency of regulation and response in the absence or presence of inducer. In general a switch system having tight regulation of the "off" state in the absence of inducer and rapid and intense response in the presence of inducer is preferred.

The ability to reversibly turn genes on and off has great utility for the analyses of gene expression and function, particularly for those genes whose products are toxic to the cell. A well characterized control mechanism in prokaryotes involves repressor proteins binding to operator DNA to prevent transcription initiation (Wray and Reznikoff (1983) *J Bacteriol* 156:1188-1191) and repressor-regulated systems have been developed for controlling expression, both in animals (Wirtz and Clayton (1995) *Science* 268:1179-1183; Deuschle et al. (1995) *Mol Cell Biol* 15:1097-1914; Furth et al. (1994) *Proc Natl Acad Sci USA* 91:9032-9306; Gossen and Bujard (1992) *Proc Natl Acad Sci USA* 89:5547-5551 and Gossen et al. (1995) *Science* 268:1766-1769) and plants (Wilde et al. (1992) *EMBO J* 11:1251-1259; Gatz et al. (1992) *Plant J* 2:397-404; Roder et al. (1994) *Mol Gen Genet.* 243-32-38 and Ulmasov et al. (1997) *Plant Mol Biol* 35:417-424).

Two major repressor based systems have been successfully exploited for regulation of plant gene expression: the lac operator-repressor system (Ulmasov et al. (1997) *Plant Mol Biol* 35:417-424; Wilde et al. (1992) *EMBO J* 11:1251-1259) and the tet operator-repressor system (Wilde et al. (1992) *EMBO J* 11:1251-1259; Gatz et al. (1992) *Plant J* 2:397-404; Roder et al. (1994) *Mol Gen Genet* 243:32-38; Ulmasov et al. (1997) *Plant Mol Biol* 35-417-424). Both are repressor/operator based-systems deriving key elements from their corresponding prokaryotic operon, namely the *E. coli* lactose operon for lac and the transposon Tn10 tetracycline operon for tet. Generally, these systems control the activity of a promoter by placing operator sequences near the transcriptional start site of a gene such that gene expression from the operon is inhibited upon the binding of the repressor protein to its cognate operator sequence. However, in the presence of an inducing agent, the binding of the repressor to its operator is inhibited, thus activating the promoter and enabling gene expression. In the lac system, isopropyl-B-D-thiogalactopyranoside (IPTG) is the commonly used inducing agent, while tetracycline and/or doxycycline are commonly used inducing agents for the tet system.

Expression of the Tn10-operon is regulated by binding of the tet repressor to its operator sequences (Beck et al. (1982) *J Bacteriol* 150:633-642; Wray and Reznikoff (1983) *J Bacteriol* 156:1188-1191). The high specificity of tetracycline repressor for the tet operator, the high efficiency of induction by tetracycline and its derivatives, the low toxicity of the inducer, as well as the ability of tetracycline to easily permeate most cells, are the basis for the application of the tet system in somatic gene regulation in eukaryotic cells from animals (Wirtz and Clayton (1995) *Science* 268:1179-1183; Gossen et al. (1995) *Science* 268:1766-1769), humans (Deuschle et al. (1995) *Mol Cell Biol* 15:1907-1914; Furth et al. (1994) *Proc Natl Acad Sci USA* 91:9302-9306; Gossen and Bujard (1992) *Proc Natl Acad Sci USA* 89:5547-5551; Gossen et al. (1995) *Science* 268:1766-1769) and plant cell cultures (Wilde et al. (1992) *EMBO J* 11:1251-1259; Gatz et al. (1992) *Plant J* 2:397-404; Roder et al. (1994) *Mol Gen Genet* 243:32-28; Ulmasov et al. (1997) *Plant Mol Biol* 35:417-424).

A number of variations of tetracycline operator/repressor systems have been devised. For example, one system based on conversion of the tet repressor to an activator was developed via fusion of the repressor to a transcriptional transactivation domain such as herpes simplex virus VP16 and the tet repressor (tTA, Gossen and Bujard (1992) *Proc Natl Acad Sci USA* 89:5547-5551). In this system, a minimal promoter is activated in the absence of tetracycline by binding of tTA to tet operator sequences, and tetracycline inactivates the transactivator and inhibits transcription. This system has been used in plants (Weinmann et al. (1994) *Plant J* 5:559-569), rat hearts (Fishman et al. (1994) *J Clin Invest* 93:1864-1868) and mice (Furth et al. (1994) *Proc Natl Acad Sci USA* 91:9302-9306). However, there were indications that the chimeric tTA fusion protein was toxic to cells at levels required for efficient gene regulation (Bohl et al. (1996) *Nat Med* 3:299-305).

Promoters modified to be regulated by tetracycline and analogs thereof are known (Matzke et al. (2003) *Plant Mol Biol Rep* 21:9-19; Padidam (2003) *Curr Op Plant Biol* 6:169-177; Gatz and Quail (1988) *Proc Natl Acad Sci USA* 85:1394-1397; Ulmasov et al. (1997) *Plant Mol Biol* 35:417-424; Weinmann, et al. (1994) *Plant J* 5:559-569). One or more tet operator sequences can be added to a promoter in order to produce a tetracycline inducible promoter. In some examples up to 7 tet operators have been introduced upstream of a minimal promoter sequence and a TetR::VP16 activation domain fusion applied in trans activates expression only in the absence of inducer (Weinmann et al. (1994) *Plant J* 5:559-569; Love et al. (2000) *Plant J* 21:579-588). A widely tested tetracycline regulated expression system for plants using the CaMV 35S promoter was developed (Gatz et al. (1992) *Plant J* 2:397-404) having three tet operators introduced near the TATA box (3XOpT 35S). The 3XOpT 35S promoter generally functioned in tobacco and potato, however toxicity and poor plant phenotype in tomato and *Arabidopsis* (Gatz (1997) *Ann Rev Plant Physiol Plant Mol Biol* 48:89-108; Corlett et al. (1996) *Plant Cell Environ* 19:447-454) were also reported. Another factor is that the tetracycline-related chemistry is rapidly degraded in the light, which tends to confine its use to testing in laboratory conditions.

TetR has been subjected to DNA shuffling to modify its inducer specificity from tetracycline to 4-de(dimethylamino)-6-deoxy-6-demethyl-tetracycline (cmt3) a related but non-inducing compound (Scholz et al. (2003) *J Mol Biol* 329:217-227) which lacks chemical side groups at positions 4 and 6 and is therefore smaller than tetracycline. The specificity of TetR was altered by narrowing the ligand binding pocket, thereby sterically blocking the natural ligand tetracycline. The starting polypeptide was a TetR(BD) chimera consisting of amino acids 1-50 from TetR(B) and residues 51-208 from TetR(D). Several rounds of evolution and selection were used to shift TetR specificity from tetracycline to cmt3. Non-inducer cmt3 had little starting activity and was brought to the level of tetracycline, yielding an improvement in activity of several thousand-fold, and tetracycline has almost no inducing activity with the mutant repressors. While the ability to shift the specificity of TetR to the cmt3 ligand is exciting, it must be kept in mind that cmt3 is highly related to the natural tetracycline ligand. Based on these experiments, it is not

obvious that TetR could be used as the basis for developing specificity to a completely different class of chemical ligands.

To produce a new chemical switch system, we re-designed the TetR system to recognize chemistry viable for use in agriculture. The re-design process was initiated by choosing a registered agrichemical compound having excellent plant uptake and distribution properties, as well as having a size and a shape reasonable for modeling into the wild type TetR ligand binding pocket. The compound chosen, thifensulfuron-methyl (Harmony®) is one of a family of commercially used sulfonylurea type herbicides inhibiting the key plant enzyme in branched chain amino acid biosynthesis, acetolactate synthase (ALS). Thifensulfuron (Ts) and related herbicides are structurally disparate to tetracycline, therefore it was unlikely they would have any starting activity with TetR. DNA shuffling is a powerful technology and can improve affinities for substrates or rate of substrate turnover by several thousand-fold, however has not yet been able to create de novo starting activity. To meet this gap in the evolution pathway a computer modeling strategy was sought that would narrow the search for meaningful amino acid diversity for shuffling. Recently developed modeling technology was used to re-train *E. coli* periplasmic binding proteins that normally bind to sugars to react to and initiate signaling with completely diverse sets of compounds such as serotonin, L-lactate and trinitrotoluene (Looger et al. (2003) *Nature* 423:185-190). Using protein design coupled with DNA shuffling and a very sensitive screening system, TetR protein variants that respond to thifensulfuron (Ts) and other related SU compounds have been identified. Following several rounds of DNA shuffling, TetR variants were developed having genetic switch capability with SU ligands (SuRs) similar to that of TetR with tetracycline inducers.

Any method of rational protein design can be used alone or in combination. For example, phylogenetic diversity within a family of protein sequences can be used to identify positions in the primary structure having amino acid substitutions, and the types of substitutions that have occurred and their impact on function. Conserved domain families can also be aligned and similarly examined to identify positions in the primary structure having amino acid substitutions and the types of substitutions that have occurred and the impact on function. The secondary structure(s) and functional domains can be evaluated and various models used to predict tolerance or impact of amino acid substitutions on structure and function. Modeling using the tertiary and/or quaternary structure and ligand, substrate and/or cofactor binding provide further insights into the effects of amino acid substitutions and/or alternate ligands, substrates and/or cofactors interactions with the polypeptide.

To examine the phylogenetic diversity of tetracycline repressors, both a broad family of tetracycline repressor proteins as well as closely related tetracycline repressors were used. Thirty-four proteins were identified and aligned to examine the amino acid diversity at various positions in the repressor family (SEQ ID NO: 1 and 402-433). The broad family of tetracycline repressors comprised a TetR(D) mutant whose structure was determined by crystallization PDB_1A6I (Orth et al. (1998) *J Mol Biol* 279:439-447) and public sequence deposit accessions A26948, AAA98409, AAD12754, AAD25094, AAD25537, AAP93923, AAR96033, AAW66496, AAW83818, ABO14708, ABS19067, CAA24908, CAC80726, CAC81917, EAY62734, NP_387455, NP_387462, NP_511232, NP_824556, P51560, YP_001220607, YP_001370475, YP_368094, YP_620166, YP_772551, ZP_00132379, ZP_01558383, and ZP_01567051. Closely related tetracycline repressors included TetR(A) P03038, TetR(B) P04483, TetR(D) P0ACT4, TetR(E) P21337 and TetR(H) P51561. The alignments of these sequences were used to look at overall sequence diversity as well as diversity in the DNA and the ligand binding domains (see, Example 1H, SEQ ID NO: 1 and 402-433).

The modular architecture of repressor proteins and the commonality of helix-turn-helix DNA binding domains allows for the creation of SuR polypeptides having altered DNA binding specificity. For example, the DNA binding specificity can be altered by fusing a SuR ligand binding domain to an alternate DNA binding domain. For example, the DNA binding domain from TetR class D can be fused to an SuR ligand binding domain to create SuR polypeptides that specifically bind to polynucleotides comprising a class D tetracycline operator. In some examples a DNA binding domain variant or derivative can be used. For example, a DNA binding domain from a TetR variant that specifically recognizes a tetO-4C operator or a tetO-6C operator could be used (Helbl and Hillen (1998) *J Mol Biol* 276:313-318; HelbI et al. (1998) *J Mol Biol* 276:319-324. The four helix bundle formed by helices α8 and α10 in both subunits can be substituted to ensure dimerization specificity when targeting two different operator specific repressor variants in the same cell to prevent heterodimerization (e.g., Rossi et al. (1998) *Nat Genet* 20:389-393; Berens and Hillen (2003) *Eur J Biochem* 270:3109-3121). In another example, the DNA binding domain from LexA repressor was fused to GAL4 wherein this hybrid protein recognized LexA operators in both *E. coli* and yeast (Brent and Ptashne (1985) *Cell* 43:729-736). In another example, all of the presumptive DNA binding or DNA-recognition R-groups of the 434 repressor were replaced by the corresponding positions of the P22 repressor. Operator binding specificity of the hybrid repressor 434R[α3(P22R)] was tested both in vivo and in vitro and each test showed that this targeted modification of 434 shifted the DNA binding specificity from 434 operator to P22 operator (Wharton and Ptashne (1985) *Nature* 316:601-605). This work was further extended by creating a heterodimer of wild type 434R and 434R[α3(P22R)] which then specifically recognized a chimeric P22/434 operator sequence (Hollis et al. (1988) *Proc Natl Acad Sci USA* 85:5834-5838). In another example, the N-terminal half of the AraC protein was fused to the LexA repressor DNA binding domain. The resulting AraC:LexA chimera dimerized, bound LexA operator, and repressed expression of a LexA operator:β-galactosidase fusion gene in an arabinose-responsive manner (Bustos and Schleif (1993) *Proc Natl Acad Sci USA* 90:5638-5642).

The isolated polynucleotides encoding SuR polypeptides can also be used as substrates for diversity-generating procedures, including mutation, recombination, and recursive recombination reactions, to produce additional SuR polynucleotide and/or polypeptide variants with desired properties. Additionally, the SuR polynucleotides can be used for diversity-generating procedures to produce polynucleotide and/or polypeptide variants having an altered characteristic as compared to the starting material, for example binding to a different ligand inducer. The diversity-generating process produces sequence alterations including single nucleotide substitutions, multiple nucleotide substitutions and insertion or deletion of regions of the nucleic acid sequence. The diversity-generating procedures can be used separately and/or in combination to produce one or more SuR variants or set of variant as well variants of encoded proteins. Individually and collectively, these procedures provide robust, widely applicable ways of generating diversified polynucleotides and polypeptides, as well as sets of polynucleotides and polypep-

tides, including, libraries. These variants and sets of variants are useful for the engineering or rapid evolution of polynucleotides, proteins, pathways, cells and/or organisms with new and/or improved characteristics. The resulting polynucleotide and/or polypeptide variants can be selected or screened for altered characteristics and/or properties, including altered ligand binding, retention of DNA binding, and/or quantification of binding properties.

Any method can be used to provide sequence diversity to a library. Many diversity-generating procedures, including multigene shuffling and methods for generating modified nucleic acid sequences are available, including for example, Soong et al. (2000) *Nat Genet* 25:436-39; Stemmer et al. (1999) *Tumor Targeting* 4:1-4; Ness et al. (1999) *Nature Biotech* 17:893-896; Chang et al. (1999) *Nature Biotech* 17:793-797; Minshull and Stemmer (1999) *Curr Op Chem Biol* 3:284-290; Christians et al. (1999) *Nature Biotech* 17:259-264; Crameri et al. (1998) *Nature* 391:288-291; Crameri et al. (1997) *Nature Biotech* 15:436-438; Zhang et al. (1997) *Proc Natl Acad Sci USA* 94:4504-4509; Patten et al. (1997) *Curr Op Biotech* 8:724-733; Crameri et al. (1996) *Nature Med* 2:100-103; Crameri et al. (1996) *Nature Biotech* 14:315-319; Gates et al. (1996) *J Mol Biol* 255:373-386; Stemmer (1996) "Sexual PCR and Assembly PCR" in The Encyclopedia of Molecular Biology (VCH Publishers, New York) pp. 447-457; Crameri and Stemmer (1995) *BioTechniques* 18:194-195; Stemmer et al. (1995) *Gene* 164:49-53; Stemmer (1995) *Science* 270:1510; Stemmer (1995) *Bio/Technology* 13:549-553; Stemmer (1994) *Nature* 370:389-391 and Stemmer (1994) *Proc Natl Acad Sci USA* 91:10747-10751. Mutational methods to generate diversity include, for example, site-directed mutagenesis (Ling et al. (1997) *Anal Biochem* 254:157-178; Dale et al. (1996) *Methods Mol Biol* 57:369-374; Smith (1985) *Ann Rev Genet* 19:423-462; Botstein and Shortle (1985) *Science* 229:1193-1201; Carter (1986) *Biochem J* 237:1-7 and Kunkel (1987) "The efficiency of oligonucleotide directed mutagenesis" in Nucleic Acids and Molecular Biology (Eckstein and Lilley, eds., Springer Verlag, Berlin). Mutagenesis methods using uracil containing templates included Kunkel (1985) *Proc Natl Acad Sci USA* 82:488-492; Kunkel et al. (1987) *Methods Enzymol* 154:367-382; and Bass et al. (1988) *Science* 242:240-245. Oligonucleotide-directed mutagenesis methods include Zoller and Smith (1983) *Methods Enzymol* 100:468-500; Zoller and Smith (1982) *Nucl Acids Res* 10:6487-6500 and Zoller and Smith (1987) *Methods Enzymol* 154:329-350. Phosphorothioate-modified DNA mutagenesis methods include Taylor et al. (1985) *Nucl Acids Res* 13:8749-8764; Taylor et al. (1985) *Nucl Acids Res* 13:8765-8787; Nakamaye and Eckstein (1986) *Nucl Acids Res* 14:9679-9698; Sayers et al. (1988) *Nucl Acids Res* 16:791-802 and Sayers et al. (1988) *Nucl Acids Res* 16:803-814. Mutagenesis methods using gapped duplex DNA include (Kramer et al. (1984) *Nucl Acids Res* 12:9441-9456; Kramer and Fritz (1987) *Methods Enzymol* 154:350-367; Kramer et al. (1988) *Nucl Acids Res* 16:7207; and Fritz et al. (1988) *Nucl Acids Res* 16:6987-6999. Additional suitable diversity-generating methods include point mismatch repair (Kramer et al. (1984) *Cell* 38:879-887); mutagenesis using repair-deficient host strains (Carter et al. (1985) *Nucl Acids Res* 13:4431-4443; and Carter (1987) *Methods Enzymol* 154:382-403); deletion mutagenesis (Eghtedarzadeh and Henikoff (1986) *Nucl Acids Res* 14:5115); restriction-selection and restriction-purification (Wells et al. (1986) *Phil Trans R Soc Lond A* 317:415-423); mutagenesis by total gene synthesis (Nambiar et al. (1984) *Science* 223:1299-1301; Sakamar and Khorana (1988) *Nucl Acids Res* 14:6361-6372; Wells et al. (1985) *Gene* 34:315-323 and Grundström et al. (1985) *Nucl Acids Res.* 13:3305-3316); double-strand break repair (Mandecki (1986) *Proc Natl Acad Sci USA* 83:7177-7181; and Arnold (1993) *Curr Op Biotech* 4:450-455). Nucleic acids can be recombined in vitro by any technique or combination of techniques including, e.g., DNAse digestion of nucleic acids to be recombined followed by ligation and/or PCR reassembly of the nucleic acids. For example, sexual PCR mutagenesis can be used in which fragmentation of the DNA molecule is followed by recombination in vitro, based on sequence similarity, between DNA molecules with different but related DNA sequences, followed by fixation of the crossover by extension in a polymerase chain reaction. Similarly, nucleic acids can be recursively recombined in vivo, e.g., by allowing recombination to occur between nucleic acids in cells. Such formats optionally provide direct recombination between nucleic acids of interest, or provide recombination between constructs, vectors, viruses, and/or plasmids comprising the nucleic acids of interest. Whole genome recombination methods can also be used wherein whole genomes of cells or other organisms are recombined, optionally including spiking of the genomic recombination mixtures with desired library components. These methods have many applications, including those in which the identity of a target gene is not known. Any of these processes can be used alone or in combination to generate polynucleotides encoding SuR polypeptides. Any of the diversity-generating methods can be used in a reiterative fashion, using one or more cycles of mutation/recombination or other diversity generation methods, optionally followed by one or more selection methods to generate additional recombinant nucleic acids.

For convenience and high throughput it will often be desirable to screen/select for desired modified nucleic acids in a microorganism, such as in a bacteria such as *E. coli*, or unicellular eukaryote such as yeast including *S. cerevisiae, S. pombe, P. pastoris* or protists such as *Chlamydomonas*, or in model cell systems such as SF9, Hela, CHO, BMS, BY2, or other cell culture systems. In some instances, screening in plant cells or plants may be desirable, including plant cell or explant culture systems or model plant systems such as *Arabidopsis*, or tobacco. In some examples throughput is increased by screening pools of host cells expressing different modified nucleic acids, either alone or as part of a gene fusion construct. Any pools showing significant activity can be deconvoluted to identify single clones expressing the desirable activity.

Recombinant constructs comprising one or more of nucleic acid sequences encoding a SuR polypeptide are provided. The constructs comprise a vector, such as, a plasmid, a cosmid, a phage, a virus, a bacterial artificial chromosome (BAC), a yeast artificial chromosome (YAC), or the like, into which a polynucleotide encoding a SuR polypeptide has been inserted. In some examples, the construct further comprises regulatory sequences, including, for example, a promoter, operably linked to the sequence. Suitable vectors are well known and include chromosomal, non-chromosomal and synthetic DNA sequences, such as derivatives of SV40; bacterial plasmids; replicons; phage DNA; baculovirus; yeast plasmids; vectors derived from combinations of plasmids and phage DNA, viral DNA such as vaccinia, adenovirus, fowl pox virus, pseudorabies, adenovirus, adeno-associated viruses, retroviruses, geminiviruses, TMV, PVX, other plant viruses, Ti plasmids, Ri plasmids and many others.

The vectors may optionally contain one or more selectable marker genes to provide a phenotypic trait for selection of transformed host cells. Usually, the selectable marker gene will encode antibiotic or herbicide resistance. Suitable genes

include those coding for resistance to the antibiotic spectinomycin or streptomycin (e.g., the aadA gene), the streptomycin phosphotransferase (SPT) gene for streptomycin resistance, the neomycin phosphotransferase (NPTII or NPTIII) gene kanamycin or geneticin resistance, the hygromycin phosphotransferase (HPT) gene for hygromycin resistance. Additional selectable marker genes include dihydrofolate reductase or neomycin resistance for eukaryotic cell culture and tetracycline or ampicillin resistance. Genes coding for resistance to herbicides include those which act to inhibit the action of glutamine synthase, such as phosphinothricin or basta (e.g., the bar gene), EPSPS, GOX, or GAT which provide resistance to glyphosate, mutant ALS (acetolactate synthase) which provides resistance to sulfonylurea type herbicides or any other known genes.

In bacterial systems a number of expression vectors are available. Such vectors include, but are not limited to, multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene); pIN vectors (Van Heeke and Schuster, (1989) *J Biol Chem* 264:5503-5509); pET vectors (Novagen, Madison Wis.) and the like. Similarly, in *S. cerevisiae* a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase and PGH may be used for production of polypeptides. For reviews, see, Ausubel and Grant et al. (1987) *Meth Enzymol* 153:516-544. A variety of expression systems can be used in mammalian host cells, including viral-based systems, such as adenovirus and rous sarcoma virus (RSV) systems. Any number of commercially or publicly available expression systems or derivatives thereof can be used.

In plant cells expression can be driven from an expression cassette integrated into a plant chromosome, or an organelle, or cytoplasmically from an episomal or viral nucleic acid. Numerous plant derived regulatory sequences have been described, including sequences which direct expression in a tissue specific manner, e.g., TobRB7, patatin B33, GRP gene promoters, the rbcS-3A promoter and the like. Alternatively, high level expression can be achieved by transiently expressing exogenous sequences of a plant viral vector, e.g., TMV, BMV, geminiviruses including WDV and the like.

Typical vectors useful for expression of nucleic acids in higher plants are known including vectors derived from the tumor-inducing (Ti) plasmid of *Agrobacterium tumefaciens* described by Rogers et al. (1987) *Meth Enzymol* 153:253-277. Exemplary *A. tumefaciens* vectors include plasmids pKYLX6 and pKYLX7 of Schardl et al. (1987) *Gene* 61:1-11 and Berger et al. (1989) *Proc Natl Acad Sci USA* 86:8402-8406 and plasmid pB101.2 that is available from Clontech Laboratories, Inc. (Palo Alto, Calif.). A variety of known plant viruses can be employed as vectors including cauliflower mosaic virus (CaMV), geminiviruses, brome mosaic virus and tobacco mosaic virus.

The SuR may be used to control expression of a polynucleotide of interest. The polynucleotide of interest may be any sequence of interest, including but not limited to sequences encoding a polypeptide, encoding an mRNA, encoding an RNAi precursor, encoding an active RNAi agent, a miRNA, an antisense polynucleotide, a ribozyme, a fusion protein, a replicating vector, a screenable marker, and the like. Expression of the polynucleotide of interest may be used to induce expression of an encoding RNA and/or polypeptide, or conversely to suppress expression of an encoded RNA, RNA target sequence, and/or polypeptide. In specific examples, the polynucleotide sequence may a polynucleotide encoding a plant hormone, plant defense protein, a nutrient transport protein, a biotic association protein, a desirable input trait, a desirable output trait, a stress resistance gene, a herbicide resistance gene, a disease/pathogen resistance gene, a male sterility, a developmental gene, a regulatory gene, a DNA repair gene, a transcriptional regulatory gene or any other polynucleotide and/or polypeptide of interest.

A number of promoters can be used in the compositions and methods. For example, a polynucleotide encoding a SuR polypeptide can be operably linked to a constitutive, tissue-preferred, inducible, developmentally, temporally and/or spatially regulated or other promoters including those from plant viruses or other pathogens which function in a plant cell. A variety of promoters useful in plants is reviewed in Potenza et al. (2004) *In Vitro Cell Dev Biol Plant* 40:1-22.

Any polynucleotide, including polynucleotides of interest, polynucleotides encoding SuRs, regulatory regions, introns, promoters, and promoters comprising TetOp sequences may be obtained and their nucleotide sequence determined, by any standard method. The polynucleotides may be chemically synthesized in their full-length or assembled from chemically synthesized oligonucleotides (Kutmeier et al. (1994) *BioTechniques* 17:242). Assembly from oligonucleotides typically involves synthesis of overlapping oligonucleotides, annealing and ligating of those oligonucleotides and PCR amplification of the ligated product. Alternatively, a polynucleotide may be isolated or generated from a suitable source including suitable source a cDNA library generated from tissue or cells, a genomic library, or directly isolated from a host by PCR amplification using specific primers to the 3' and 5' ends of the sequence or by cloning using an nucleotide probe specific for the polynucleotide of interest. Amplified nucleic acid molecules generated by PCR may then be cloned into replicable cloning vectors using standard methods. The polynucleotide may be further manipulated using any standard methods including recombinant DNA techniques, vector construction, mutagenesis and PCR (see, e.g., Sambrook et al. (1990) Molecular Cloning, A Laboratory Manual, 2d Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.; Ausubel et al., Eds. (1998) Current Protocols in Molecular Biology, John Wiley and Sons, NY).

Any method for introducing a sequence into a cell or organism can be used, as long as the polynucleotide or polypeptide gains access to the interior of at least one cell. Methods for introducing sequences into plants are known and include, but are not limited to, stable transformation, transient transformation, virus-mediated methods, and sexual breeding. Stably incorporated indicates that the introduced polynucleotide is integrated into a genome and is capable of being inherited by progeny. Transient transformation indicates that an introduced sequence does not integrate into a genome such that it is heritable by progeny from the host. Any means can be used to bring together a SuR and polynucleotide of interest operably linked to a promoter comprising TetOp including, for example, stable transformation, transient delivery, cell fusion, sexual crossing or any combination thereof.

Transformation protocols as well as protocols for introducing polypeptides or polynucleotide sequences into plants may vary depending on the type of plant or plant cell targeted for transformation. Suitable methods of introducing polypeptides and polynucleotides into plant cells include microinjection (Crossway et al. (1986) *Biotechniques* 4:320-334 and U.S. Pat. No. 6,300,543), electroporation (Riggs et al. (1986) *Proc Natl Acad Sci USA* 83:5602-5606, *Agrobacterium*-mediated transformation (U.S. Pat. Nos. 5,563,055 and 5,981,840), direct gene transfer (Paszkowski et al. (1984) *EMBO J* 3:2717-2722), ballistic particle acceleration (U.S. Pat. Nos. 4,945,050, 5,879,918, 5,886,244 and 5,932,782; Tomes et al. (1995) in Plant Cell, Tissue and Organ Culture: Fundamental Methods, ed. Gamborg and Phillips (Springer-Verlag, Ber-

lin); McCabe et al. (1988) *Biotechnology* 6:923-926). Also see, Weissinger et al. (1988) *Ann Rev Genet* 22:421-477; Sanford et al. (1987) Particulate *Science and Technology* 5:27-37; Christou et al. (1988) *Plant Physiol* 87:671-674; Finer and McMullen (1991) *In Vitro Cell Dev Biol* 27P:175-182 (soybean); Singh et al. (1998) *Theor Appl Genet* 96:319-324; Datta et al. (1990) *Biotechnology* 8:736-740; Klein et al. (1988) *Proc Natl Acad Sci USA* 85:4305-4309; Klein et al. (1988) *Biotechnology* 6:559-563; U.S. Pat. Nos. 5,240,855, 5,322,783 and 5,324,646; Klein et al. (1988) *Plant Physiol* 91:440-444; Fromm et al. (1990) *Biotechnology* 8:833-839; Hooykaas-Van Slogteren et al. (1984) *Nature* 311:763-764; U.S. Pat. No. 5,736,369; Bytebier et al. (1987) *Proc Natl Acad Sci USA* 84:5345-5349; De Wet et al. (1985) in The Experimental Manipulation of Ovule Tissues, ed. Chapman et al. (Longman, New York), pp. 197-209; Kaeppler et al. (1990) *Plant Cell Rep* 9:415-418; Kaeppler et al. (1992) *Theor Appl Genet* 84:560-566; D'Halluin et al. (1992) *Plant Cell* 4:1495-1505; Li et al. (1993) *Plant Cell Rep* 12:250-255; Christou and Ford (1995) *Ann Bot* 75:407-413 and Osjoda et al. (1996) *Nat Biotechnol* 14:745-750. Alternatively, polynucleotides may be introduced into plants by contacting plants with a virus, or viral nucleic acids. Methods for introducing polynucleotides into plants via viral DNA or RNA molecules are known, see, e.g., U.S. Pat. Nos. 5,889,191, 5,889,190, 5,866,785, 5,589,367, 5,316,931 and Porta et al. (1996) *Mol Biotech* 5:209-221.

The term plant includes plant cells, plant protoplasts, plant cell tissue cultures from which a plant can be regenerated, plant calli, plant clumps and plant cells that are intact in plants or parts of plants such as embryos, pollen, ovules, seeds, leaves, flowers, branches, fruit, kernels, ears, cobs, husks, stalks, roots, root tips, anthers and the like. Progeny, variants and mutants of the regenerated plants are also included.

In some examples, a SuR may be introduced into a plastid, either by transformation of the plastid or by directing a SuR transcript or polypeptide into the plastid. Any method of transformation, nuclear or plastid, can be used, depending on the desired product and/or use. Plastid transformation provides advantages including high transgene expression, control of transgene expression, ability to express polycistronic messages, site-specific integration via homologous recombination, absence of transgene silencing and position effects, control of transgene transmission via uniparental plastid gene inheritance and sequestration of expressed polypeptides in the organelle which can obviate possible adverse impacts on cytoplasmic components (e.g., see, reviews including Heifetz (2000) *Biochimie* 82:655-666; Daniell et al. (2002) *Trends Plant Sci* 7:84-91; Maliga (2002) *Curr Op Plant Biol* 5:164-172; Maliga (2004) *Ann Rev Plant Biol* 55-289-313; Daniell et al. (2005) *Trends Biotechnol* 23:238-245 and Verma and Daniell (2007) *Plant Physiol* 145:1129-1143).

Methods and compositions of plastid transformation are well known, for example, transformation methods include (Boynton et al. (1988) *Science* 240:1534-1538; Svab et al. (1990) *Proc Natl Acad Sci USA* 87:8526-8530; Svab et al. (1990) *Plant Mol Biol* 14:197-205; Svab et al. (1993) *Proc Natl Acad Sci USA* 90:913-917; Golds et al. (1993) *Bio/Technology* 11:95-97; O'Neill et al. (1993) *Plant J* 3:729-738; Koop et al. (1996) *Planta* 199:193-201; Kofer et al. (1998) *In Vitro Plant* 34:303-309; Knoblauch et al. (1999) *Nat Biotechnol* 17:906-909); as well as plastid transformation vectors, elements, and selection (Newman et al. (1990) *Genetics* 126:875-888; Goldschmidt-Clermont, (1991) *Nucl Acids Res* 19:4083-4089; Carrer et al. (1993) *Mol Gen Genet* 241:49-56; Svab et al. (1993) *Proc Natl Acad Sci USA* 90:913-917; Verma and Daniell (2007) *Plant Physiol* 145:1129-1143).

Methods and compositions for controlling gene expression in plastids are well known including (McBride et al. (1994) *Proc Natl Acad Sci USA* 91:7301-7305; Lössl et al. (2005) *Plant Cell Physiol* 46:1462-1471; Heifetz (2000) *Biochemie* 82:655-666; Surzycki et al. (2007) *Proc Natl Acad Sci USA* 104:17548-17553; U.S. Pat. Nos. 5,576,198 and 5,925,806; WO 2005/0544478), as well as methods and compositions to import polynucleotides and/or polypeptides into a plastid, including translational fusion to a transit peptide (e.g., Comai et al. (1988) *J Biol Chem* 263:15104-15109).

The SuR polynucleotides and polypeptides provide a means for regulating plastid gene expression via a chemical inducer that readily enters the cell. For example, using the T7 expression system for chloroplasts (McBride et al. (1994) *Proc Natl Acad Sci USA* 91:7301-7305) the SuR could be used to control nuclear T7 polymerase expression. Alternatively, an SuR-regulated promoter could be integrated into the plastid genome and operably linked to the polynucleotide(s) of interest and the SuR expressed and imported from the nuclear genome, or integrated into the plastid. In all cases, application of a sulfonylurea compound is used to efficiently regulate the polynucleotide(s) of interest.

Any type of cell and/or organism, prokaryotic or eukaryotic, can be used with the SuR methods and compositions. For example, any bacterial cell system can be transformed with the compositions. For example, methods of *E. coli, Agrobacterium* and other bacterial cell transformation, plasmid preparation and the use of phages are detailed, for example, in Current Protocols in Molecular Biology (Ausubel, et al., (eds.) (1994) a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc.).

The SuR systems can be used with any eukaryotic cell line, including yeasts, protists, algae, insect cells, avian or mammalian cells. For example, many commercially and/or publicly available strains of *S. cerevisiae* are available, as are the plasmids used to transform these cells. For example, strains are available from the American Type Culture Collection (ATCC, Manassas, Va.) and include the Yeast Genetic Stock Center inventory, which moved to the ATCC in 1998. Other yeast lines, such as *S. pombe* and *P. pastoris*, and the like are also available. For example, methods of yeast transformation, plasmid preparation, and the like are detailed, for example, in Current Protocols in Molecular Biology (Ausubel et al. (eds.) (1994) a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., see Unit 13 in particular). Transformation methods for yeast include spheroplast transformation, electroporation, and lithium acetate methods. A versatile, high efficiency transformation method for yeast is described by Gietz and Woods ((2002) *Methods Enzymol* 350:87-96) using lithium acetate, PEG 3500 and carrier DNA.

The SuRs can be used in mammalian cells, such as CHO, HeLa, BALB/c, fibroblasts, mouse embryonic stem cells and the like. Many commercially available competent cell lines and plasmids are well known and readily available, for example from the ATCC (Manassas, Va.). Isolated polynucleotides for transformation and transformation of mammalian cells can be done by any method known in the art. For example, methods of mammalian and other eukaryotic cell transformation, plasmid preparation, and the use of viruses are detailed, for example, in Current Protocols in Molecular Biology (Ausubel et al. (eds.) (1994) a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., see, Unit 9 in particular). For example, many methods

are available, such as calcium phosphate transfection, electroporation, DEAE-dextran transfection, liposome-mediated transfection, microinjection as well as viral techniques.

Any plant species can be used with the SuR methods and compositions, including, but not limited to, monocots and dicots. Examples of plants include, but are not limited to, corn (*Zea mays*), *Brassica* spp. (e.g., *B. napus, B. rapa, B. juncea*), castor, palm, alfalfa (*Medicago sativa*), rice (*Oryza sativa*), rye (*Secale cereale*), sorghum (*Sorghum bicolor, Sorghum vulgare*), millet (e.g., pearl millet (*Pennisetum glaucum*), proso millet (*Panicum miliaceum*), foxtail millet (*Setaria italica*), finger millet (*Eleusine coracana*)), sunflower (*Helianthus annuus*), safflower (*Carthamus tinctorius*), wheat (*Triticum aestivum*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), potato (*Solanum tuberosum*), peanuts (*Arachis hypogaea*), cotton (*Gossypium barbadense, Gossypium hirsutum*), sweet potato (*Ipomoea batatus*), cassava (*Manihot esculenta*), coffee (*Coffea* spp.), coconut (*Cocos nucifera*), pineapple (*Ananas comosus*), citrus trees (*Citrus* spp.), cocoa (*Theobroma cacao*), tea (*Camellia sinensis*), banana (*Musa* spp.), avocado (*Persea americana*), fig (*Ficus casica*), guava (*Psidium guajava*), mango (*Mangifera indica*), olive (*Olea europaea*), papaya (*Carica papaya*), cashew (*Anacardium occidentale*), *macadamia* (*Macadamia integrifolia*), almond (*Prunus amygdalus*), sugar beets (*Beta vulgaris*), sugarcane (*Saccharum* spp.), *Arabidopsis thaliana*, oats (*Avena* spp.), barley (*Hordeum* spp.), leguminous plants such as guar beans, locust bean, fenugreek, garden beans, cowpea, mungbean, fava bean, lentils, and chickpea, vegetables, ornamentals, grasses and conifers. Vegetables include tomatoes (*Lycopersicon esculentum*), lettuce (e.g., *Lactuca sativa*), green beans (*Phaseolus vulgaris*), lima beans (*Phaseolus limensis*), peas (*Pisium* spp., *Lathyrus* spp.), and *Cucumis* species such as cucumber (*C. sativus*), cantaloupe (*C. cantalupensis*), and musk melon (*C. melo*). Ornamentals include azalea (*Rhododendron* spp.), hydrangea (*Macrophylla hydrangea*), hibiscus (*Hibiscus rosasanensis*), roses (*Rosa* spp.), tulips (*Tulipa* spp.), daffodils (*Narcissus* spp.), petunias (*Petunia hybrida*), carnation (*Dianthus caryophyllus*), poinsettia (*Euphorbia pulcherrima*), and chrysanthemum. Conifers include pines, for example, loblolly pine (*Pinus taeda*), slash pine (*Pinus elliotii*), ponderosa pine (*Pinus ponderosa*), lodgepole pine (*Pinus contorta*), and Monterey pine (*Pinus radiata*), Douglas fir (*Pseudotsuga menziesii*); Western hemlock (*Tsuga canadensis*), Sitka spruce (*Picea glauca*), redwood (*Sequoia sempervirens*), true firs such as silver fir (*Abies amabilis*) and balsam fir (*Abies balsamea*) and cedars such as Western red cedar (*Thuja plicata*) and Alaska yellow cedar (*Chamaecyparis nootkatensis*).

The plant cells and/or tissue that have been transformed may be grown into plants using conventional methods (see, e.g., McCormick et al. (1986) *Plant Cell Rep* 5:81-84). These plants may then be grown and self-pollinated, backcrossed, and/or outcrossed, and the resulting progeny having the desired characteristic identified. Two or more generations may be grown to ensure that the characteristic is stably maintained and inherited and then seeds harvested. In this manner transformed/transgenic seed having a DNA construct comprising a polynucleotide of interest and/or modified polynucleotide encoding an SuR stably incorporated into their genome are provided. A plant and/or a seed having stably incorporated the DNA construct can be further characterized for expression, agronomics and copy number.

Sequence identity may be used to compare the primary structure of two polynucleotides or polypeptide sequences, describe the primary structure of a first sequence relative to a second sequence, and/or describe sequence relationships such as variants and homologues. Sequence identity measures the residues in the two sequences that are the same when aligned for maximum correspondence. Sequence relationships can be analyzed using computer-implemented algorithms. The sequence relationship between two or more polynucleotides or two or more polypeptides can be determined by computing the best alignment of the sequences and scoring the matches and the gaps in the alignment, which yields the percent sequence identity and the percent sequence similarity. Polynucleotide relationships can also be described based on a comparison of the polypeptides each encodes. Many programs and algorithms for comparison and analysis of sequences are known. Unless otherwise stated, sequence identity/similarity values provided herein refer to the value obtained using GAP Version 10 (GCG, Accelrys, San Diego, Calif.) using the following parameters: % identity and % similarity for a nucleotide sequence using GAP Weight of 50 and Length Weight of 3, and the nwsgapdna.cmp scoring matrix; % identity and % similarity for an amino acid sequence using GAP Weight of 8 and Length Weight of 2, and the BLOSUM62 scoring matrix (Henikoff and Henikoff (1992) *Proc Natl Acad Sci USA* 89:10915-10919). GAP uses the algorithm of Needleman and Wunsch (1970) *J Mol Biol* 48:443-453, to find the alignment of two complete sequences that maximizes the number of matches and minimizes the number of gaps.

Alternatively, polynucleotides and/or polypeptides can be evaluated using other sequence tools. For example, polynucleotides and/or polypeptides can be evaluated using a BLAST alignment tool. A local alignment gaps consists simply of a pair of sequence segments, one from each of the sequences being compared. A modification of Smith-Waterman or Sellers algorithms will find all segment pairs whose scores cannot be improved by extension or trimming, called high-scoring segment pairs (HSPs). The results of the BLAST alignments include statistical measures to indicate the likelihood that the BLAST score can be expected from chance alone. The raw score, S, is calculated from the number of gaps and substitutions associated with each aligned sequence wherein higher similarity scores indicate a more significant alignment. Substitution scores are given by a look-up table (see PAM, BLOSUM). Gap scores are typically calculated as the sum of G, the gap opening penalty and L, the gap extension penalty. For a gap of length n, the gap cost would be G+Ln. The choice of gap costs, G and L is empirical, but it is customary to choose a high value for G (10-15) and a low value for L (1-2). The bit score, S', is derived from the raw alignment score S in which the statistical properties of the scoring system used have been taken into account. Bit scores are normalized with respect to the scoring system, therefore they can be used to compare alignment scores from different searches. The E-Value, or expected value, describes the likelihood that a sequence with a similar score will occur in the database by chance. It is a prediction of the number of different alignments with scores equivalent to or better than S that are expected to occur in a database search by chance. The smaller the E-Value, the more significant the alignment. For example, an alignment having an E value of e^{-117} means that a sequence with a similar score is very unlikely to occur simply by chance. Additionally, the expected score for aligning a random pair of amino acid is required to be negative, otherwise long alignments would tend to have high score independently of whether the segments aligned were related. Additionally, the BLAST algorithm uses an appropriate substitution matrix, nucleotide or amino acid and for gapped alignments uses gap creation and extension penalties. For example, BLAST alignment and comparison of polypeptide

sequences are typically done using the BLOSUM62 matrix, a gap existence penalty of 11 and a gap extension penalty of 1. Unless otherwise stated, scores reported from BLAST analyses were done using the BLOSUM62 matrix, a gap existence penalty of 11 and a gap extension penalty of 1.

UniProt protein sequence database is a repository for functional and structural protein data and provides a stable, comprehensive, fully classified, richly and accurately annotated protein sequence knowledgebase, with extensive cross-references and querying interfaces freely accessible to the scientific community. The UniProt site has a tool, UniRef, that provides a cluster of proteins have 50%, 90% or 100% sequence identity to a protein sequence of interest from the database. For example, using TetR(B) (UniProt reference P04483) gives a cluster of 18 proteins having 90% sequence identity to P04483:

RefID	Protein Name	Species	Length
P04483	TetR class B from transposon Tn10	*E. coli*	207
B1VCF0	TetR protein	*E. coli*	208
A0ZSZ1	Tetracycline resistant gene repressor	*Photobacterium* sp. TC21	208
A4LA82	Tetracycline repressor protein	*Edwardsiella tarda*	208
A4V9K4	Tetracycline repressor	*Salmonella enterica*	208
A8R6K3	Tetracycline repressor protein	*Salmonella enterica* subsp. *enterica* serovar *Choleraesuis*	208
Q573N4	Tetracycline repressor protein	uncultured bacterium	208
Q7BQ37	TetR	*Shigella flexneri*	208
Q9S455	TetR	*Salmonella typhi*	208
A4IUI5	Tetracycline repressor protein R, class B	*Yersinia ruckeri*	207
Q1A2K5	Tetracycline resistance repressor protein	*E. coli*	207
Q6MXH5	TetR class B from transposon tn10	*Serratia marcescens*	207
Q79VX4	TetR protein	*Salmonella typhimurium*	207
Q7AZW7	Tet repressor protein	*Pasteurella aerogenes*	207
Q7AK84	Repressor of tet operon	Plasmid R100	207
Q6QR72	Tetracycline repressor protein	*E. coli*	208
Q93F26	Tet repressor	*Shigella flexneri* 2a	208
Q8L0M9	Putative tetracycline repressor protein	*Neisseria meningitidis*	205

These protein sequences can be used as sources for sequence diversity for protein design and/or directed evolution of the ligand binding domain. Further, these protein sequences can be used as sources for operator binding domains for chimeric repressor proteins, or for design and/or evolution of the operator binding domain.

The properties, domains, motifs and function of tetracycline repressors are well known, as are standard techniques and assays to evaluate any derived repressor comprising one or more amino acid substitutions. The structure of the class D TetR protein comprises 10 alpha helices with connecting loops and turns. The 3 N-terminal helices form the DNA-binding HTH domain, which has an inverse orientation as compared to HTH motifs in other DNA-binding proteins. The core of the protein, formed by helices 5-10, comprises the dimerization interface domain, and for each monomer comprises the binding pocket for ligand/effector and divalent cation cofactor (Kisker et al. (1995) *J Mol Biol* 247:260-180; Orth et al. (2000) *Nat Struct Biol* 7:215-219). Any amino acid change may comprise a non-conservative or conservative amino acid substitution. Conservative substitutions generally refer to exchanging one amino acid with another having similar chemical and/or structural properties (see, e.g., Dayhoff et al. (1978) Atlas of Protein Sequence and Structure, Natl Biomed Res Found, Washington, D.C.). Different clustering of amino acids by similarity have been developed depending on the property evaluated, such as acidic vs. basic, polar vs. non-polar, amphipathic and the like and be used when evaluating the possible effect of any substitution or combination of substitutions.

Numerous variants of TetR have been identified and/or derived and extensively studied. In the context of the tetracycline repressor system, the effects of various mutations, modifications and/or combinations thereof have been used to extensively characterize and/or modify the properties of tetracycline repressors, such as cofactor binding, ligand binding constants, kinetics and dissociation constants, operator binding sequence constraints, cooperativity, binding constants, kinetics and dissociation constants and fusion protein activities and properties. Variants include TetR variants with a reverse phenotype of binding the operator sequence in the presence of tetracycline or an analog thereof, variants having altered operator binding properties, variants having altered operator sequence specificity and variants having altered ligand specificity and fusion proteins. See, for example, Isackson and Bertrand (1985) *Proc Natl Acad Sci USA* 82:6226-6230; Smith and Bertrand (1988) *J Mol Biol* 203: 949-959; Altschmied et al. (1988) *EMBO J* 7:4011-4017; Wissmann et al. (1991) *EMBO J* 10:4145-4152; Baumeister et al. (1992) *J Mol Biol* 226:1257-1270; Baumeister et al. (1992) *Proteins* 14:168-177; Gossen and Bujard (1992) *Proc Natl Acad Sci USA* 89:5547-5551; Wasylewski et al. (1996) *J Protein Chem* 15:45-58; Berens et al. (1997) *J Biol Chem* 272:6936-6942; Baron et al. (1997) *Nucl Acids Res* 25:2723-2729; Helbl and Hillen (1998) *J Mol Biol* 276:313-318; Urlinger et al. (2000) *Proc Natl Acad Sci USA* 97:7963-7968; Kamionka et al. (2004) *Nucl Acids Res* 32:842-847; Bertram et al. (2004) *J Mol Microbiol Biotechnol* 8:104-110; Scholz et al. (2003) *J Mol Biol* 329: 217-227; and patent publication US 2003/0186281.

The three-dimensional structures of tetracycline repressors, and tetracycline repressor variants, coupled to ligand and/or co-factor(s), and bound to operator sequence are known (see, for example, Kisker et al. (1995) *J Mol Biol* 247:260-280; Orth et al. (1998) *J Mol Biol* 279:439-447; Orth et al. (1999) *Biochemistry* 38:191-198; Orth et al. (2000) *Nat Struct Biol* 7:215-219; Luckner et al. (2007) *J Mol Biol* 368: 780-790) providing extremely well characterized structure(s), identification of domains and individual amino acids associated with various functions and binding properties, and predictive model(s) for the potential effects of any amino acid substitution(s), as well as the possible structural bases for the phenotype(s) of known tetracycline repressor mutants. One example of percent sequence identity observed within tetracycline repressor family members is shown below.

% polypeptide sequence identity between TetR family members

TetR Class	A (P03038)	E (P21337)	B (P04483)	D (P0ACT4)	H (P51561)
A (P03038)	100	44	51	48	50
E (P21337)		100	51	49	50
B (P04483)			100	64	64
D (P0ACT4)				100	58
H (P51561)					100

EXAMPLES

Example 1

Evolution of TetR for Recognition by Sulfonylurea Compounds

A. Computational Modeling

The 3-D crystal structures of the class D tetracycline repressor (isolated from *E. coli*; TET-bound dimer, 1DU7 (Orth et al. (2000) *Nat Struct Biol* 7:215-219); and DNA-bound dimer, 1QPI (Orth et al. (2000) *Nat Struct Biol* 7:215-219)), were used as the design scaffold for computational replacement of the tetracycline (TET) molecule by the thifensulfuron-methyl (Ts, Harmony®) molecule in the ligand binding pocket. TET and sulfonylureas (SUs) are generally similar in size and have aromatic ring-based structures with hydrogen bond donors and acceptors, potentially allowing SU binding to a mutated TetR. However, there are notable differences between the tetracycline family and SU family of molecules. TET is internally rigid and fairly flat, with one highly-hydrogen-bonding face with hydroxyls and ketones, logP ~−0.3. Sulfonylureas (SUs) are more highly flexible and aromatic, with a core sulfonyl-urea moiety typically connecting a substituted benzene, pyridine, or thiophene (as in the case of Harmony®) on one side with a substituted pyrimidine or 1,3,5-triazine on the other side. Although having different functional groups, the logP of Harmony® is similar (~0.02 at pH 7) to that of tet. A best-posed Harmony® molecule was positioned by molecular modeling in the TetR binding pocket in silico (FIG. **1**). Based on this model, seventeen amino acid residue positions (60, 64, 82, 86, 100, 104, 105, 113, 116, 134, 135, 138 and 139 from monomer A and positions 147, 151, 174 and 177 from monomer B, using TetR(B) numbering) were determined to be in sufficiently close proximity to a docked Harmony® as to be recruited into a binding surface. Computational side-chain optimization was employed to design sets of amino acids at each of the 17 positions deemed to be most compatible with SU binding. This resulted in a library with (4, 5, 4, 4, 5, 3, 8, 11, 10, 10, 8, 8, 7, 9, 6, 7 and 5) amino acids at the 17 positions, for a total designed library size of 4×10^{13}. The choice of amino acids at the library positions was dictated by steric and physicochemical considerations to fit ligand docking into the ligand pocket.

The wild type class B TetR from Tn10 was chosen as the starting molecule for generation of shuffling derivatives (SEQ ID NO: 2). It is slightly different than the sequence used in computational design (P0ACT4, class D, for which the high-resolution crystal structure 1DU7 is available), but only subtly affects ligand binding. A comparison of TetR(D) (SEQ ID NO: 401) and TetR(B) (SEQ ID NO: 2) is shown below with positions involved in tet recognition and binding in bold:

```
          1       .       .       .       .       .
1DU7        SRLNRESVIDAALELLNETGIDGLTTRKLAQKLGIEQPTLYWHVKNKRALLDALAVEI
Class B   MGSRLDKSKVINSALELLNEVGIEGLTTRKLAQKLGVEQPTLYWHVKNKRALLDALAIEM

         61       .       .       .       .       .       .
1DU7      LARHHDYSLPAAGESWQSFLRNNAMSFRRALLRYRDGAKVHLGTRPDEKQYDTVETQLRF
Class B   LDRHHTHFCPLEGESWQDFLRNNAKSFRCALLSHRDGAKVHLGTRPTEKQYETLENQLAF

        121       .       .       .       .       .       .
1DU7      MTENGFSLRDGLYAISAVSHFTLGAVLEQQEHTAALTDRPAAPDENLPPLLREALQIMDS
Class B   LCQQGFSLENALYALSAVGHFTLGCVLEDQEHQVAKEERETPTTDSMPPLLRQAIELFDH

        181       .       .       208
1DU7      DDGEQAFLHGLESLIRGFEVQLTALLQIV
Class B   QGAEPAFLFGLELIICGLEKQLKCESGS-
```

The starting polynucleotide used to express TetR was synthesized commercially and restriction sites were added for functionality in library construction and further manipulation (DNA2.0, Menlo Park, Calif., USA). Added restriction sites include an NcoI site at the 5' end, a SacI site 5' of the ligand binding domain (LBD) and an AscI site following the stop codon. This allows library construction to be localized in a ~480 bp DNA segment containing the ligand binding region to avoid inadvertent mutations in the other regions, such as the DNA binding domain. The synthetic gene was operably linked downstream of an arabinose inducible promoter, P_{BAD}, using Ncoi/AscI to create TetR expression vector pVER7314 (FIG. **2**). The addition of the NcoI site at the 5' end of the coding region resulted in the insertion of a glycine after the N-terminal methionine at amino acid position one (SEQ ID NO: 2). This sequence was used as the wild type TetR control in all assays unless otherwise noted, and observed activity was equivalent to TetR without the serine insertion (SEQ ID NO: 1). However, all references to amino acid positions and changes designed and observed use the amino acid numbering of wild type TetR(B) (207 aa) e.g., SEQ ID NO: 1.

B. Library Design

Due to the large number of designed substitutions at many positions in close proximity with one another the computed library (Table 1, Designed Library) was not easily encodable with a small number of degenerate codons. This is particularly evident in sequence regions such as amino acids 134, 135, 138 and 139, which could reasonably be encoded by a single primer. For this reason, the sequence library fabricated and tested in the lab featured the designed amino acid set at 6/17 positions, slightly enlarged at 1/17 positions, and fully degenerate (NNK codon) at 10/17 positions (Table 1). This resulted in much higher predicted sequence diversity, a total of 3×10^{19} sequences.

TABLE 1

Resi-due	WT resi-due	Designed Library	Actual Library
60	L	A L K M	A L K M
64	H	A N Q H L	A N Q H L
82	N	A N S T	A N S T
86	F	M F W Y	M F W Y
100	H	H M F W Y	All 20 aa's
104	R	A R G	A R G
105	P	A N D G P S T V	All 20 aa's
113	L	A R N D Q E K M S T V	A R N D Q E K M S T V I P L G H
116	Q	A R N Q E I K M T V	All 20 aa's
134	L	A R I L K M F W Y V	All 20 aa's

TABLE 1-continued

Resi-due	WT resi-due	Designed Library	Actual Library
135	S	A R N Q H K S T	A R N Q H K S T
138	G	A H K M F S Y W	All 20 aa's
139	H	A R Q H L K Y	All 20 aa's
147	E	A R Q E H L K M Y	All 20 aa's
151	H	A Q H K I L	All 20 aa's
174	I	A R Q E L K M	All 20 aa's
177	F	A R L K M	All 20 aa's

The constructed library, termed 'L1', was encoded with a total of fifty oligonucleotides (Table 2) rather than the thousands that would have been required to completely specify the designed target library. Table 2 also includes two PCR amplification primers.

TABLE 2

Oligo ID	Oligo Sequence	SEQ ID
L1:01	TATTGGCATGTAAAAAATAAGCGAGCTCTGCTCGACGCCTTA	833
L1:02	GCCATTGAGATGAWGGATAGGCACCWGACTCACTTTGCCCT	834
L1:03	GCCATTGAGATGAWGGATAGGCACMATACTCACTTTGCCCT	835
L1:04	GCCATTGAGATGAWGGATAGGCACGCGACTCACTTTGCCCT	836
L1:05	GCCATTGAGATGACGGATAGGCACCWGACTCACTTTGCCCT	837
L1:06	GCCATTGAGATGACGGATAGGCACMATACTCACTTTGCCCT	838
L1:07	GCCATTGAGATGACGGATAGGCACGCGACTCACTTTGCCCT	839
L1:08	GCCATTGAGATGATGGATAGGCACCWGACTCACTTTGCCCT	840
L1:09	GCCATTGAGATGATGGATAGGCACMATACTCACTTTGCCCT	841
L1:10	GCCATTGAGATGATGGATAGGCACGCGACTCACTTTGCCCT	842
L1:11	TTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAACGCT	843
L1:12	TTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATDCTGCT	844
L1:13	AAAAGTTWTAGATGTGCTTTACTAAGTCATCGCGATGGAGCA	845
L1:14	AAAAGTTGGAGATGTGCTTTACTAAGTCATCGCGATGGAGCA	846
L1:15	AAAAGTATGAGATGTGCTTTACTAAGTCATCGCGATGGAGCA	847
L1:16	AAAGTANNKTTAGGTACAGCGNNKACAGAAAAACAGTATGAA	848
L1:17	AAAGTANNKTTAGGTACASGCNNKACAGAAAAACAGTATGAA	849
L1:18	ACTVNSGAAAATNNKTTAGCCTTTTTATGCCAACAAGGTTTT	850
L1:19	TCACTAGAGAATGCATTATATGCANNSRCCGCTGTGNNKNNK	851
L1:20	TCACTAGAGAATGCATTATATGCANNSMGCGCTGTGNNKNNK	852
L1:21	TCACTAGAGAATGCATTATATGCANNSMAKGCTGTGNNKNNK	853
L1:22	TTTACTTTAGGTTGCGTATTGNNKGATCAAGAGNNKCAAGTC	854
L1:23	GCTAAAGAAGAAAGGGAAACACCTACTACTGATAGTATGCCG	855
L1:24	CCATTATTACGACAAGCTNNKGAATTANNKGATCACCAAGGT	856

TABLE 2-continued

Oligo ID	Oligo Sequence	SEQ ID
L1:25	GCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATATGC	857
L1:26	GGATTAGAAAAACAACTTAAATGTGAAAGTGGGTCTTAAGGC	858
L1:27	CCTATCCWTCATCTCAATGGCTAAGGCGTCGAGCAGAGCTCG	859
L1:28	CCTATCCGCCATCTCAATGGCTAAGGCGTCGAGCAGAGCTCG	860
L1:29	CCTATCCAGCATCTCAATGGCTAAGGCGTCGAGCAGAGCTCG	861
L1:30	TTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGTCWGGTG	862
L1:31	TTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGTATKGTG	863
L1:32	TTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGTCGCGTG	864
L1:33	TAAAGCACATCTAWAACTTTTAGCGTTATTACGTAAAAAATC	865
L1:34	TAAAGCACATCTCCAACTTTTAGCGTTATTACGTAAAAAATC	866
L1:35	TAAAGCACATCTCATACTTTTAGCGTTATTACGTAAAAAATC	867
L1:36	TAAAGCACATCTAWAACTTTTAGCAGHATTACGTAAAAAATC	868
L1:37	TAAAGCACATCTCCAACTTTTAGCAGHATTACGTAAAAAATC	869
L1:38	TAAAGCACATCTCATACTTTTAGCAGHATTACGTAAAAAATC	870
L1:39	CGCTGTACCTAAMNNTACTTTTGCTCCATCGCGATGACTTAG	871
L1:40	GCSTGTACCTAAMNNTACTTTTGCTCCATCGCGATGACTTAG	872
L1:41	GGCTAAMNNATTTTCSNBAGTTTCATACTGTTTTTCTGTMNN	873
L1:42	ATATAATGCATTCTCTAGTGAAAAACCTTGTTGGCATAAAAA	874
L1:43	CAATACGCAACCTAAAGTAAAMNNMNNCACAGCGGYSNNTGC	875
L1:44	CAATACGCAACCTAAAGTAAAMNNMNNCACAGCGCKSNNTGC	876
L1:45	CAATACGCAACCTAAAGTAAAMNNMNNCACAGCMTKSNNTGC	877
L1:46	TGTTTCCCTTTCTTCTTTAGCGACTTGMNNCTCTTGATCMNN	878
L1:47	MNNAGCTTGTCGTAATAATGGCGGCATACTATCAGTAGTAGG	879
L1:48	GAATAAGAAGGCTGGCTCTGCACCTTGGTGATCMNNTAATTC	880
L1:49	TTTAAGTTGTTTTTCTAATCCGCATATGATCAATTCAAGGCC	881
L1:50	GGGAACTTCGGCGCGCCTTAAGACCCACTTTCACA	882
L1:5'	CATGTAAAAAATAAGCGAGCTCTG	883
L1:3'	GGGAACTTCGGCGCGCCTTAAGAC	884

Assembly of the 'L1' oligos was carried out by overlap extension (Ness, et al., (2002) *Nat Biotech* 20:1251-1255) to generate a PCR fragment bordered by SacI/AscI restriction sites. Conditions for assembly of all library fragments were as follows: oligonucleotides representing the library are normalized to a concentration of 10 μM and then equal volumes mixed to create a 10 μM pool. PCR amplification of library fragments was performed in six identical 25 μl reactions containing: 1 M pooled library oligos; 0.5 μM of each rescue primer: L1:5' and L1:3' and 200 M dNTP's in a Herculase II directed reaction (Stratagene, La Jolla, Calif., USA). Conditions for PCR were 98° C. for 1 min (initial denature), followed by 25 cycles of 95° C. denature for 20 seconds, annealing for 45 seconds between 45° C. and 55° C. (gradient), then extending the template for 30 seconds at 72° C. A final extension of 72° C. for 5 minutes completes the reaction. Wild type TetR(B) is excised from the P_{BAD}-tetR expression vector pVER7314 by digestion with SacI/AscI. The pVER7314 backbone fragment is treated with calf intestinal phosphatase and purified, then the fully extended library fragment pool (~500 bp) digested with SacI/AscI restriction enzymes are inserted to generate the L1 plasmid library. Approximately 50 random clones from library L1 were sequenced and the information compiled for quality control purposes. The results indicated that nearly all amino acids targeted in the diversity set were represented (data not shown). Sequencing revealed that 17% of the sequences contained stop codons. This is less than the predicted 27% (e.g., 10 positions having 1/32 codons be a stop codon, $1-(31/32)^{10}$~27%). Additionally, sequence analysis showed that 13% of the clones had frame shifts due

to mistakes in the overlap extension process. Thus, overall approximately 30% of the library consisted of clones encoding truncated polypeptides.

C. Screen Set Up

In order to test the library for rare clones reacting to thifensulfuron-methyl (Ts) a sensitive *E. coli* based genetic screen was developed. The screen is a modification of an established assay system (Wissmann et al. (1991) *Genetics* 128:225-232). The screen consists of two parts: a repressor pre-screen followed by an induction screen. For this purpose an *E. coli* strain was developed having both functionalities. For the repressor prescreen a genetic cascade was developed whereby an nptIII gene encoding kanamycin resistance is under the control of a lac promoter. The lac promoter is repressed by the Lac repressor encoded by lacI, whose expression is in turn controlled by the tet promoter (PtetR). The tet promoter is repressed by TetR which blocks LacI production and thus ultimately enables kanamycin resistance to be expressed.

Since the tet regulon has bivalent promoters, one promoter for tetR and one promoter for tetA, the same strain was engineered with the *E. coli* lacZ gene encoding enzyme reporter β-galactosidase under control of the tetA promoter (PtetA). The dual regulon encoding both lacI and lacZ was then bordered by strong transcriptional terminators: the *E. coli* RNA ribosomal operon terminator rrnB T1-T2 (Ghosh et al. (1991) *J Mol Biol* 222:59-66) and the *E. coli* RNA polymerase subunit C terminator rpoC, such that spurious transcripts read in the direction of either tet promoter would not interfere with expression of any other transcript. In the presence of functional TetR, the strain exhibits a lac^- phenotype and colonies can be easily scored for induction by novel chemistry with X-gal, wherein induction gives increased blue colony color. In addition, induction with novel chemistry in liquid cultures can be measured quantitatively by employing β-galactosidase enzyme assays with either colorimetric or fluorimetric substrates.

A further refinement of the host strain is that the to/C locus was knocked out with the incoming Plac-nptIII reporter. This was done to obtain better penetration of SU compounds into *E. coli* (Robert LaRossa—DuPont: personal communication). A strong transcriptional terminator, T22 from *S. typhimurium* phage P22, was placed upstream of the lac promoter to prevent unregulated leaky expression of the conditional kanamycin resistance marker. The name of the final engineered strain is *E. coli* KM3.

The population of shuffled tetR LBD's was cloned into an Ap^r/ColE1 based vector pVER7314 behind the P_{BAD} promoter. This was designed to enable fine control of TetR expression by variation of arabinose concentrations in the growth medium (Guzman et al. (1995) *J Bacteriol* 177:4121-4130). Despite being under the control of the P_{BAD} promoter, TetR protein is expressed at a sufficient level in the absence of added arabinose to enable selection for kanamycin resistance in strain KM3. Nevertheless, expression can be increased by addition of arabinose, for example, if a change in assay stringency is desired.

D. Library Screening

Following assembly of L1 oligos and capture in vector pVER7314, the resulting library was transformed into *E. coli* strain KM3 and plated on LB containing 50 μg/ml carbenicillin to select for library plasmids, and 60 μg/ml kanamycin to select for the active repressor population in the absence of target ligand ("apo-repressors"). DNA sequence analysis of this selected population indicated that this step highly enriched several library positions, suggesting that few amino acid combinations in the ligand binding domain lead to a conformation compatible with DNA binding by the N-terminal domain. In addition, this step eliminated clones with premature stop codons and or frame shift mutations. Subsequently, these apo-repressor sequences were screened for alteration in repressor activity in the presence of Harmony® (Ts). This was done by replica plating the Km^r pre-selected population from liquid cultures in 384-well format onto M9 agar containing 0.1% glycerol as carbon source, 0.04% casamino acids (to prevent branched chain amino acid starvation caused by sulfonylurea application), 50 μg/ml carbenicillin for plasmid maintenance, 0.004% X-gal to detect galactosidase activity, and +/−SU inducer Ts at 20 μg/ml. Initial hits were identified from a population of nearly 20,000 colonies screened for response to Ts following incubation at 30° C. for 2 days. Fourteen putative 'hits' identified were then re-tested under the same conditions but in 96-well format (FIG. **3**).

DNA sequence analysis revealed that clones L1-3 and L1-19 are identical and that the most intensely responding hits (L-2, -3(19), -5, -9, -11 and -20) had significant enrichment at several library positions, indicating an involvement in ligand interaction, directly or indirectly. The same library was then re-screened to identify a further 10 hits to bring the total number of clones to 23.

All 23 putative hits were subsequently screened in the same plate assay format with a panel of nine sulfonylurea (SU) compounds registered for commercial use (Table 3), wherein 11 hits were found to respond significantly to other SU ligands (Table 4). For this experiment, *E. coli* clones encoding L1 hits or wt TetR (SEQ ID NO: 2) were arrayed in 96-well format and stamped onto M9 X-gal assay media with or without test SU compounds at 20 μg/ml. Following 48 hrs growth at 30° C. the plates were digitally imaged and the colony color intensity converted to relative values of β-galactosidase activity. Inducers used: thifensulfuron (Ts), metsulfuron (Ms), sulfometuron (Sm), ethametsulfuron (Es), tribenuron (Tb), chlorimuron (Ci), nicosulfuron (Ns), rimsulfuron (Rs), chlorsulfuron (Cs) at 20 ppm and anhydrotetracycline (atc) as the positive control at 0.4 μM for induction of wt TetR. Surprisingly, some sulfonylurea compounds, particularly chlorimuron, ethametsulfuron, and chlorsulfuron were more potent activators than the starting ligand Harmony®.

TABLE 3

SU Compound		
Common Name	Product Name	Commercial Use
Thifensulfuron-methyl (Ts)	Harmony ®	Cereals, corn, soybean
Metsulfuron-methyl (Ms)	Ally ®	Cereals, pasture
Sulfometuron-methyl (Sm)	Oust ®	Vegetation management
Ethametsulfuron-methyl (Es)	Muster ®	Canola
Tribenuron-methyl (Tb)	Express ®	Cereal, sunflower
Chlorimuron-ethyl (Ci)	Classic ®	Soybean
Nicosulfuron (Ns)	Accent ®	Corn
Rimsulfuron (Rs)	Matrix ®	Corn, tomato, potato
Chlorsulfuron (Cs)	Glean ®	Cereals

TABLE 4

	Inducer										
clone	None	Ts	Ms	Sm	Es	Tb	Ci	Ns	Rs	Cs	atc
L1-2	1.0	1.6	1.9	4.7	5.8	1.7	13.6	1.3	1.3	4.1	1.2
L1-7	0.0	0.1	0.2	6.4	0.1	0.2	16.5	0.1	0.2	3.1	0.0
L1-9	0.3	1.1	1.2	0.6	11.8	0.4	9.8	0.3	0.4	23.6	0.3
L1-20	1.4	2.6	12.4	6.0	15.0	2.6	13.5	1.6	2.0	22.0	2.0
L1-22	0.1	0.0	0.1	17.2	0.3	0.3	10.4	0.2	0.1	0.2	0.0
L1-24	0.1	0.3	0.4	3.1	0.2	1.6	22.1	0.3	0.3	3.3	0.1
L1-28	0.0	0.1	18.8	1.1	0.8	0.3	14.6	0.1	0.2	5.8	0.0
L1-29	0.0	0.0	13.5	2.7	1.7	0.3	20.9	0.1	0.1	15.8	0.0
L1-31	0.3	0.9	0.5	0.9	13.7	0.1	1.1	0.5	0.4	1.4	0.4
L1-38	9.5	16.7	14.7	18.3	14.8	15.8	15.3	8.7	9.5	14.0	6.4
L1-44	0.2	1.9	2.9	0.4	2.4	0.4	6.7	0.4	0.3	12.0	0.2
TetR	0.0	0.0	0.0	0.1	0.0	0.1	0.1	0.1	0.1	0.0	25.0

The amino acid substitutions for these eleven top hits are summarized in Table 5. The sequences are compared to wild type TetR(B) and only positions having differences are shown using the numbering according to TetR(B) (e.g., SEQ ID NO: 1). A dash in the alignment indicates no change relative to wt TetR ligand binding domain.

TABLE 5

	60	64	82	86	100	104	105	113	116	134	135	138	139	147	151	164	174	177	203	205
TetR(B)	L	H	N	F	H	R	P	L	Q	L	S	G	H	E	H	D	I	F	C	S
L1-02	–	A	–	–	C	A	V	M	I	R	R	I	A	W	Y	–	R	I	–	–
L1-07	–	N	–	W	V	A	I	H	P	I	A	R	R	V	R	–	S	Q	–	R
L1-09	–	A	–	M	C	G	F	A	S	M	Q	C	I	L	L	–	L	K	–	–
L1-20	M	Q	–	M	F	A	W	V	L	–	N	A	T	W	K	–	H	G	S	–
L1-22	M	–	T	Y	C	A	I	K	N	R	Q	R	V	F	M	–	S	L	S	–
L1-24	–	N	S	M	L	A	V	T	S	I	R	R	T	V	R	–	K	L	–	–
L1-28	–	A	–	M	W	A	W	P	V	S	R	V	T	T	K	–	W	L	–	–
L1-29	M	Q	T	M	W	–	W	P	M	W	–	C	N	S	R	–	W	S	–	–
L1-31	–	A	–	M	M	–	A	V	R	V	R	H	W	I	M	–	Y	Y	–	–
L1-38	A	–	–	M	W	A	W	T	M	W	R	T	M	R	W	–	L	G	–	–
L1-44	–	A	–	Y	Y	A	V	A	–	V	K	A	G	W	S	A	V	A	–	–

The initial screenings of library 1 also detected library members having reverse repressor activity (SEQ ID NO: 1206-1213), wherein the polypeptide was bound to the operator in the presence of SU ligand. These hits showed β-galactosidase expression without SU ligand, which was substantially reduced upon addition of the ligand, for example thifensulfuron. These hits were subsequently screened in the same plate assay format as described above with the panel of nine sulfonylurea (SU) compounds registered for commercial use (Table 3), wherein 8 hits were found to respond significantly to other SU ligands (Table 6).

TABLE 6

	Inducer										
clone	Blank	Ts	Ms	Sm	Es	Tb	Ci	Ns	Rs	Cs	atc
L1-18	1.34	1.13	0.79	0.94	0.37	1.65	0.36	1.44	2.55	1.22	2.35
L1-21	2.88	0.79	0.89	2.39	0.61	2.13	0.07	2.74	2.31	0.89	2.81
L1-25	1.17	0.64	0.32	0.63	0.13	1.72	0.11	1.21	1.08	0.28	1.22
L1-33	7.59	5.51	4.29	5.02	2.11	4.71	0.76	5.34	10.32	3.74	8.25
L1-34	2.37	2.97	1.47	2.00	1.33	2.26	0.43	2.91	2.30	0.85	3.68
L1-36	1.52	0.48	0.38	0.50	0.20	0.57	0.21	1.81	1.84	0.24	1.70
L1-39	3.65	1.42	0.75	0.91	0.60	0.97	0.49	3.03	4.72	0.89	4.92

TABLE 6-continued

clone	Inducer										
	Blank	Ts	Ms	Sm	Es	Tb	Ci	Ns	Rs	Cs	atc
L1-41	4.05	1.46	0.56	0.67	0.18	1.41	0.39	2.75	4.05	0.61	4.21
TetR	0.00	0.08	0.08	0.23	0.06	0.13	0.18	0.18	0.20	0.15	10.45

The amino acid substitutions for these eight reverse repressor hits (SEQ ID NO: 1206-1213 encoded by SEQ ID NO: 1214-1221) are summarized in Table 7. The sequences are compared to wild type TetR(B) and only positions having differences are shown using the numbering according to TetR (B) (e.g., SEQ ID NO: 1). A dash in the alignment indicates no change relative to wt TetR ligand binding domain.

-continued

Position	Amino Acid Bias Observed		
	Apo repressor	SU Induction	Both
113	A (not G, P)	A, I (not D, G)	A
116	(not GL)	M, V (not A, R)	

TABLE 7

Clone	60	63	64	82	84	86	100	104	105	113	116	121	134	135	138	139	147	151	163	174	177	205	206
TetR wt	L	H	H	N	K	F	H	R	P	L	Q	C	L	S	G	H	E	H	T	I	F	S	G
SU-TetR-18	–	–	L	–	–	M	W	G	F	K	R	–	I	R	S	R	Q	P	–	V	–	–	E
SU-TetR-21	–	–	A	–	–	–	C	A	G	–	C	–	R	–	V	C	F	M	–	A	L	–	–
SU-TetR-25	–	–	A	T	–	M	L	A	T	–	L	Y	W	Q	W	R	I	T	–	V	K	T	–
SU-TetR-33	–	–	A	N	–	M	Q	A	A	–	K	–	H	–	T	Q	R	G	–	R	R	–	–
SU-TetR-34	A	–	N	A	R	M	Y	A	R	T	V	–	V	R	P	R	L	R	–	R	T	–	–
SU-TetR-36	A	–	–	A	–	M	R	A	W	H	V	–	–	K	S	G	K	M	–	T	V	–	–
SU-TetR-39	M	–	Q	T	–	Y	I	–	W	R	V	–	W	A	–	P	R	R	–	L	M	–	–
SU-TetR-41	–	N	Q	–	–	W	M	–	N	A	G	–	C	R	W	T	D	T	S	M	K	–	–

E. Correlation of First Round Shuffling Results with the Structural Model

Significant enrichment occurred at most library positions, where enrichment includes biases favoring particular amino acids and biases disfavoring particular amino acids. The initial screening involved two stages to identify both repressor and de-repressor functions. Enrichment occurred in both stages of screening. In the first stage, positions were enriched by the selection for "apo repressors", that is, proteins that repress gene transcription in the absence of ligand. In the second stage, positions were enriched by the selection for "activators", that is, proteins that allow gene transcription in the presence of ligand. Positions may be enriched by either selection criterion, by both criteria, or by neither. The first-round screening results for repressor activity are summarized below:

Position	Amino Acid Bias Observed		
	Apo repressor	SU Induction	Both
60		L (not K)	
64		Q, N (not L, A)	
82	N (not A, T)	A (not N, S)	
86	(not M)	M (not W)	
100	R (not K, Q)	C, W (not H, K, Q)	
104	G	A	
105	C, G, L, V (not H, K)	L, W (not G, S)	L

-continued

Position	Amino Acid Bias Observed		
	Apo repressor	SU Induction	Both
134	M, S	I, R, W (not G)	
135	K, R (not H, S)	Q, R (not A, T)	R
138	(not T)	A, C, R, V (not L, P, Q, T)	
139	R (not H)	T (not L, P)	
147	(not A, C)	R, W (not A, S)	
151	R (not C, G, Q)	M, R (not V)	R
174	V (not L, R)	W (not F, L)	
177	T (not S)	K, L (not P, T)	

Enrichment at the activator level was consistent with the computational model of SU binding: sterically-selected positions (e.g., 60, 86, 104, 105, 113, 138, 139) occurred in closest proximity to the modeled ligand, electrostatically-selected positions (e.g., 135, 147, 151, 177) occurred near the modeled sulfonyl moiety, and aromatically-selected positions (e.g., 100, 105, 147, 174) were modeled to form aromatic stacking interactions with the planar ring structures in thifensulfuron. Enrichment at the apo repressor level was consistent with the predicted mode of DNA binding: enriched positions were modeled to be capable of modulating association of the repressor protein with the DNA operator sequence.

In the case of the SU induction screen, enrichment was evidenced by both over-represented amino acids that were modeled to form favorable interactions with the ligand (e.g., methionine (M) and valine (V) at position 116 were modeled

to pack well against the triazine ring of the SU molecule) and by under-represented amino acids that were modeled to produce unfavorable interactions with the ligand (e.g., tryptophan ('W') at position 86 was modeled to be too large to accommodate ligand in binding pocket). In the case of the apo repressor, enrichment took the form both of over-represented amino acids that were modeled to give rise to a functional repressor conformation capable of binding the DNA operator sequence in the absence of ligand (e.g., alanine ('A') at position 113, which maintains the structural integrity of this α-helix) and of under-represented amino acids that were modeled to disrupt the actively repressing conformation in the absence of ligand (e.g., glycine and proline ('P') at position 113, which reduce the structural integrity of this α-helix).

Results from different rounds of screening and selection may produce altered enrichments at some positions, as the result of interactions with other amino acids selected, or with the small molecule. Enriched sequences only demonstrate that side-chains can contribute to active inducers, and do not rule out any amino acids. Thus, selected hits are likely to represent only a subset of possible active sequences. A variety of possible ligand-binding modes and protein interactions may be possible for the same SU molecule, and thus enrichment of several types of side-chains at a specific position may represent multiple populations of active proteins. Computational modeling of the enriched sequences is useful insofar as it enables the prioritization of amino acid diversity for rounds of screening and selection.

Altogether, these enrichment results support the overall computational model and facilitated additional design. Several positions which were constructed as fully-degenerate codons (all 20 amino acids) returned first-round screening results consistent with the suggested computational model. For example, computational modeling suggested that the aromatic side-chains W, Y and F at position 100 would favor SU binding. The first-round library was screened with a degenerate codon at this position, and the amino acids W, Y and F were significantly enriched. Designed libraries allow sequence diversity to be narrowed at library positions, with an emphasis on decreasing diversity at coupled positions such that fully degenerate codons may be avoided. Additionally, more positions could be recruited for diversification to achieve greater coverage of a higher-quality, more focused sequence library. This allows construction of a library with sufficiently few members to screen with good coverage, yet sufficient diversity to discover sequences with detectable activity. Such sequences may then be improved by introducing more diversity at library positions, with a screen or selection choosing optimal clones, independent of model predictions. In this way, the combination of computational modeling and directed evolution allows the generation of engineered sequences unlikely to be discovered by either technique separately.

F. Second-round Shuffling

The original library was designed to thifensulfuron, but once induction activity was established with other SU compounds having potentially better soil and in planta stability properties than the original ligand, the evolution process was re-directed towards these alternative ligands. Of particular interest were herbicides metsulfuron, sulfometuron, ethametsulfuron and chlorsulfuron. For this objective, parental clones L1-9, -22, -29 and -44 were chosen for further shuffling. Clone L1-9 has strong activity on both ethametsulfuron and chlorsulfuron; clone L1-22 has strong sulfometuron activity; clone L1-29 has moderate metsulfuron activity; and clone L1-44 has moderate activity on metsulfuron, ethametsulfuron and chlorsulfuron (Table 4). No clones found in the initial screen were exceptionally reactive to metsulfuron. These four clones were also chosen due to their relatively strong repressor activity, showing low β-gal background activity without inducer. Strong repressor activity is important for establishing a system which is both highly sensitive to the presence od inducer, and tightly off in the absence of inducer.

Based on the sequence information from parental clones L1-9, -22, -29 and -44, two second round libraries were designed, constructed and screened. The first library, L2, consisted of a 'family' shuffle whereby the amino acid diversity between the selected parental clones was varied using synthetic assembly of oligonucleotides to find clones improved in responsiveness to any of the four new target ligands. A summary of the diversity used and the resulting hit sequences for library L2 is shown in Table 8.

TABLE 8

	Amino acid residue position									
Clone	60	64	82	86	100	104	105	113	116	134
wt	L	H	N	F	H	R	P	L	Q	L
Parents										
L1-9	–	A	–	M	C	G	F	A	S	M
L1-22	M	–	T	Y	C	A	I	K	N	R
L1-29	M	Q	T	M	W	–	W	P	M	W
L1-44	–	A	–	Y	Y	A	V	A	–	V
Hits										
L2-2	–	Q	–	M	C	–	F	K	–	V
L2-9	M	Q	–	M	Y	–	W	A	–	W
L2-10	–	A	–	M	W	G	W	K	M	M
L2-13	–	Q	–	M	C	–	W	A	–	W

TABLE 8-continued

	Amino acid residue position									
L2-14	**M**	**A**	**–**	**M**	**C**	**–**	**W**	**A**	**M**	**V**
L2-18	**M**	**Q**	**T**	**M**	**W**	**–**	**W**	**A**	**–**	**M**
L1-45	A	Q	–	W	W	G	L	P	V	T
Unselected frequency	random	random	random	random	W > C, Y	R >> G, A	W > V > I, F	random	random	random

Clone wt	135 S	138 G	139 H	147 E	151 H	164 D	174 I	177 F	203 C	Inducer preference atc
Parents										
L1-9	Q	C	I	L	L	–	L	K	–	4, 9 (weak)
L1-22	Q	R	V	F	M	–	S	L	S	3
L1-29	–	C	N	S	R	–	W	S	–	9 (weak)
L1-44	K	A	G	W	S	A	V	A	–	9 (weak)
Hits										
L2-2	–	R	I	W	M	–	W	L	–	4 (inverse)
L2-9	–	A	I	W	S	–	S	K	–	9 (leaky)
L2-10	–	R	I	L	L	–	W	K	–	4 (leaky)
L2-13	Q	R	I	S	M	–	V	K	–	9
L2-14	**–**	**R**	**V**	**F**	**S**	**A**	**L**	**K**	**–**	9
L2-18	**–**	**R**	**N**	**F**	**L**	**A**	**W**	**K**	**–**	9
L1-45	Q	R	–	G	R	–	A	L	–	3, 4
Unselected frequency	S >> Q, K	A >> C, R	H, N > V > I	random	random	random	random	random	C >> S	

The oligonucleotides used to construct the library are shown in Table 9. The L2 oligonucleotides were assembled, cloned and screened as per the protocol described for library L1 except that each ligand was tested at 2 ppm to increase the stringency of the assay, which is a 10-fold reduction from 1st round library screening concentration.

TABLE 9

Oligo	Sequence	SEQ ID
L2:01	TATTGGCATGTAAAAAATAAGCGAGCTCTGCTCGACGCCTTA	885
L2:02	GCCATTGAGATGWTGGATAGGCACCASACTCACTTTTGCCCT	886
L2:03	GCCATTGAGATGWTGGATAGGCACGCAACTCACTTTTGCCCT	887
L2:04	TTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAMTGCT	888
L2:05	AAAAGTTACAGATGTGCTTTACTAAGTCATCGCGATGGAGCA	889
L2:06	AAAAGTATGAGATGTGCTTTACTAAGTCATCGCGATGGAGCA	890
L2:07	AAAGTATRTTTAGGTACACGCDTCACAGAAAAACAGTATGAA	891
L2:08	AAAGTATRTTTAGGTACACGCTGGACAGAAAAACAGTATGAA	892
L2:09	AAAGTATRTTTAGGTACAGSTDTCACAGAAAAACAGTATGAA	893
L2:10	AAAGTATRTTTAGGTACAGSTTGGACAGAAAAACAGTATGAA	894
L2:11	AAAGTATGGTTAGGTACACGCDTCACAGAAAAACAGTATGAA	895
L2:12	AAAGTATGGTTAGGTACACGCTGGACAGAAAAACAGTATGAA	896

TABLE 9-continued

Oligo	Sequence	SEQ ID
L2:13	AAAGTATGGTTAGGTACAGSTDTCACAGAAAAACAGTATGAA	897
L2:14	AAAGTATGGTTAGGTACAGSTTGGACAGAAAAACAGTATGAA	898
L2:15	ACTAAAGAAAATARCTTAGCCTTTTTATGCCAACAAGGTTTT	899
L2:16	ACTAAAGAAAATCAATTAGCCTTTTTATGCCAACAAGGTTTT	900
L2:17	ACTAAAGAAAATATGTTAGCCTTTTTATGCCAACAAGGTTTT	901
L2:18	ACTSCTGAAAATARCTTAGCCTTTTTATGCCAACAAGGTTTT	902
L2:19	ACTSCTGAAAATCAATTAGCCTTTTTATGCCAACAAGGTTTT	903
L2:20	ACTSCTGAAAATATGTTAGCCTTTTTATGCCAACAAGGTTTT	904
L2:21	TCACTAGAGAATGCATTATATGCARTGAGTGCTGTGGCTAWT	905
L2:22	TCACTAGAGAATGCATTATATGCARTGAGTGCTGTGGCTGKT	906
L2:23	TCACTAGAGAATGCATTATATGCARTGAGTGCTGTGYGCAWT	907
L2:24	TCACTAGAGAATGCATTATATGCARTGAGTGCTGTGYGCGKT	908
L2:25	TCACTAGAGAATGCATTATATGCARTGMAAGCTGTGGCTAWT	909
L2:26	TCACTAGAGAATGCATTATATGCARTGMAAGCTGTGGCTGKT	910
L2:27	TCACTAGAGAATGCATTATATGCARTGMAAGCTGTGYGCAWT	911
L2:28	TCACTAGAGAATGCATTATATGCARTGMAAGCTGTGYGCGKT	912
L2:29	TCACTAGAGAATGCATTATATGCAWGGAGTGCTGTGGCTAWT	913
L2:30	TCACTAGAGAATGCATTATATGCAWGGAGTGCTGTGGCTGKT	914
L2:31	TCACTAGAGAATGCATTATATGCAWGGAGTGCTGTGYGCAWT	915
L2:32	TCACTAGAGAATGCATTATATGCAWGGAGTGCTGTGYGCGKT	916
L2:33	TCACTAGAGAATGCATTATATGCAWGGMAAGCTGTGGCTAWT	917
L2:34	TCACTAGAGAATGCATTATATGCAWGGMAAGCTGTGGCTGKT	918
L2:35	TCACTAGAGAATGCATTATATGCAWGGMAAGCTGTGYGCAWT	919
L2:36	TCACTAGAGAATGCATTATATGCAWGGMAAGCTGTGYGCGKT	920
L2:37	TTTACTTTAGGTTGCGTATTGTKGGATCAAGAGAGMCAAGTC	921
L2:38	TTTACTTTAGGTTGCGTATTGTKGGATCAAGAGMTGCAAGTC	922
L2:39	TTTACTTTAGGTTGCGTATTGTYTGATCAAGAGAGMCAAGTC	923
L2:40	TTTACTTTAGGTTGCGTATTGTYTGATCAAGAGMTGCAAGTC	924
L2:41	GCTAAAGAAGAAAGGGAAACACCTACTACTGMTAGTATGCCG	925
L2:42	CCATTATTACGACAAGCTAGTGAATTATTGGATCACCAAGGT	926
L2:43	CCATTATTACGACAAGCTAGTGAATTAKCAGATCACCAAGGT	927
L2:44	CCATTATTACGACAAGCTAGTGAATTAAAGGATCACCAAGGT	928
L2:45	CCATTATTACGACAAGCTTKGGAATTATTGGATCACCAAGGT	929
L2:46	CCATTATTACGACAAGCTTKGGAATTAKCAGATCACCAAGGT	930
L2:47	CCATTATTACGACAAGCTTKGGAATTAAAGGATCACCAAGGT	931
L2:48	CCATTATTACGACAAGCTGTAGAATTATTGGATCACCAAGGT	932
L2:49	CCATTATTACGACAAGCTGTAGAATTAKCAGATCACCAAGGT	933
L2:50	CCATTATTACGACAAGCTGTAGAATTAAAGGATCACCAAGGT	934
L2:51	GCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATATGC	935

TABLE 9-continued

Oligo	Sequence	SEQ ID
L2:52	GGATTAGAAAAACAACTTAAATSCGAAAGTGGGTCTTAA	936
L2:53	CCTATCCAWCATCTCAATGGCTAAGGCGTCGAGCAGAGCTCG	937
L2:54	TTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGTSTGGTG	938
L2:55	TTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGTTGCGTG	939
L2:56	TAAAGCACATCTGTAACTTTTAGCAKTATTACGTAAAAAATC	940
L2:57	TAAAGCACATCTCATACTTTTAGCAKTATTACGTAAAAAATC	941
L2:58	GCGTGTACCTAAAYATACTTTTGCTCCATCGCGATGACTTAG	942
L2:59	ASCTGTACCTAAAYATACTTTTGCTCCATCGCGATGACTTAG	943
L2:60	GCGTGTACCTAACCATACTTTTGCTCCATCGCGATGACTTAG	944
L2:61	ASCTGTACCTAACCATACTTTTGCTCCATCGCGATGACTTAG	945
L2:62	GGCTAAGYTATTTTCTTTAGTTTCATACTGTTTTTCTGTGAH	946
L2:63	GGCTAATTGATTTTCTTTAGTTTCATACTGTTTTTCTGTGAH	947
L2:64	GGCTAACATATTTTCTTTAGTTTCATACTGTTTTTCTGTGAH	948
L2:65	GGCTAAGYTATTTTCAGSAGTTTCATACTGTTTTTCTGTGAH	949
L2:66	GGCTAATTGATTTTCAGSAGTTTCATACTGTTTTTCTGTGAH	950
L2:67	GGCTAACATATTTTCAGSAGTTTCATACTGTTTTTCTGTGAH	951
L2:68	GGCTAAGYTATTTTCTTTAGTTTCATACTGTTTTTCTGTCCA	952
L2:69	GGCTAATTGATTTTCTTTAGTTTCATACTGTTTTTCTGTCCA	953
L2:70	GGCTAACATATTTTCTTTAGTTTCATACTGTTTTTCTGTCCA	954
L2:71	GGCTAAGYTATTTTCAGSAGTTTCATACTGTTTTTCTGTCCA	955
L2:72	GGCTAATTGATTTTCAGSAGTTTCATACTGTTTTTCTGTCCA	956
L2:73	GGCTAACATATTTTCAGSAGTTTCATACTGTTTTTCTGTCCA	957
L2:74	ATATAATGCATTCTCTAGTGAAAAACCTTGTTGGCATAAAAA	958
L2:75	CAATACGCAACCTAAAGTAAAAWTAGCCACAGCACTCAYTGC	959
L2:76	CAATACGCAACCTAAAGTAAAAMCAGCCACAGCACTCAYTGC	960
L2:77	CAATACGCAACCTAAAGTAAAAWTGCRCACAGCACTCAYTGC	961
L2:78	CAATACGCAACCTAAAGTAAAAMCGCRCACAGCACTCAYTGC	962
L2:79	CAATACGCAACCTAAAGTAAAAWTAGCCACAGCTTKCAYTGC	963
L2:80	CAATACGCAACCTAAAGTAAAAMCAGCCACAGCTTKCAYTGC	964
L2:81	CAATACGCAACCTAAAGTAAAAWTGCRCACAGCTTKCAYTGC	965
L2:82	CAATACGCAACCTAAAGTAAAAMCGCRCACAGCTTKCAYTGC	966
L2:83	CAATACGCAACCTAAAGTAAAAWTAGCCACAGCACTCCWTGC	967
L2:84	CAATACGCAACCTAAAGTAAAAMCAGCCACAGCACTCCWTGC	968
L2:85	CAATACGCAACCTAAAGTAAAAWTGCRCACAGCACTCCWTGC	969
L2:86	CAATACGCAACCTAAAGTAAAAMCGCRCACAGCACTCCWTGC	970
L2:87	CAATACGCAACCTAAAGTAAAAWTAGCCACAGCTTKCCWTGC	971
L2:88	CAATACGCAACCTAAAGTAAAAMCAGCCACAGCTTKCCWTGC	972
L2:89	CAATACGCAACCTAAAGTAAAAWTGCRCACAGCTTKCCWTGC	973
L2:90	CAATACGCAACCTAAAGTAAAAMCGCRCACAGCTTKCCWTGC	974

TABLE 9-continued

Oligo	Sequence	SEQ ID
L2:91	TGTTTCCCTTTCTTCTTTAGCGACTTGKCTCTCTTGATCCMA	975
L2:92	TGTTTCCCTTTCTTCTTTAGCGACTTGCAKCTCTTGATCCMA	976
L2:93	TGTTTCCCTTTCTTCTTTAGCGACTTGKCTCTCTTGATCARA	977
L2:94	TGTTTCCCTTTCTTCTTTAGCGACTTGCAKCTCTTGATCARA	978
L2:95	ACTAGCTTGTCGTAATAATGGCGGCATACTAKCAGTAGTAGG	979
L2:96	CMAAGCTTGTCGTAATAATGGCGGCATACTAKCAGTAGTAGG	980
L2:97	TACAGCTTGTCGTAATAATGGCGGCATACTAKCAGTAGTAGG	981
L2:98	GAATAAGAAGGCTGGCTCTGCACCTTGGTGATCCAATAATTC	982
L2:99	GAATAAGAAGGCTGGCTCTGCACCTTGGTGATCTGMTAATTC	983
L2:100	GAATAAGAAGGCTGGCTCTGCACCTTGGTGATCCTTTAATTC	984
L2:101	TTTAAGTTGTTTTTCTAATCCGCATATGATCAATTCAAGGCC	985
L2:102	GGGAACTTCGGCGCGCCTTAAGACCCACTTTCGSA	986

G. Screening Results for Library L2

Nearly 8,000 colonies arising from the repressor prescreen were subjected to the activation screen on M9 assay medium containing potential inducers ethametsulfuron, sulfometuron, metsulfuron or chlorsulfuron at 2 ppm. After 48 hours of incubation at 30° C. the plates were analyzed. Twelve putative hits from these plates were re-arrayed into 96-well format and retested on the same set of media (Table 10). Only clone L2-14 had a strong induction response and tight regulation to any of the inducers at this lower 2 ppm dose, wherein it had a strong response to Cs and low background without inducer. Clone L2-18 had a moderate response to this ligand and low background. Clone L2-9 also responded well to Cs, but had higher background activity without inducer. No isolates showed a significant response to metsulfuron. An unexpected observation was that parent clone L1-9 has sensitivity to 2 ppm Es. Sequence analysis of the hit clones from library 2 (Table 6) indicates that F86M, G138R and F177K were preferred substitutions in the responding hits. This is especially striking at position 138 where A is far over represented in the unselected population and yet five clones have R at this position while only one has alanine. R105W could also be important, but in the random sequence population W105 was already biased over C or Y.

TABLE 10

		Inducer				
	Sample	No inducer	Ms	Sm	Es	Cs
L1 parents	L1-9	0.9	0.9	0.9	14.8	2.2
	L1-22	0.2	0.2	1.8	0.2	0.2
	L1-29	0.1	0.3	0.2	0.3	0.2
	L1-44	0.6	0.5	0.6	1.0	1.9
L2 hits	L2-9	2.2	4.2	6.3	5.0	14.0
	L2-10	1.3	1.1	1.7	7.7	1.7
	L2-13	1.8	1.9	2.4	2.0	8.4
	L2-14	0.5	1.0	0.9	0.9	17.0
	L2-18	0.3	0.1	0.2	0.3	4.6
	TetR	0.1	0.0	0.1	0.1	0.1

H. Second Round Library L4 Assembly

Another second round library, L4, was created from synthetic oligonucleotides using clone L1-9 as the template and injecting novel amino acid diversity into the sequence based on phylogenetic comparison of 34 TetR and related molecules. A multiple sequence alignment of SEQ ID NO: 1 and SEQ ID NO: 402-433 generated using GCG SeqWeb PILEUP (Accelrys, San Diego, Calif.) under default parameters of gap weight=8, gap length weight=2, and the BLOSUM62 matrix is shown below:

```
                1                                                  50
ZP_01558383     ~~~mkdtg.a rltrdtvmra aldllnevgi dglstrrlae rlgvqsptly

YP_772551       ~~~mkdts.t rltrdtvmra aldllnevgi dglstrrlae rlgvqsptly

YP_620166       ~~~mkdtg.t rltrdtvlra alnlldevgi dglstrrlae rlgvqsptly

EAY62734        miemkdtg.a rltrdtvlra alnlldevgi dglstrrlae rlgvqsptly

YP_368094       ~~~mkdtg.a rltrdtvlra alelldevgi dglstrklae rlgvqsptly

AAP93923        ~mseknta.a rltretvlrg alalldd igi dalstrrlaq hlgvqsptly

AAW66496        ~mskkdiapq rltreivlrt aldmlneegi dsittrklaq rlgiksptly

CAA24908        ~~~~~~~~mt klqpntvira aldllnevgv dglttrklae rlgvqqpaly

P03038          ~~~~~~~~mt klqpntvira aldllnevgv dglttrklae rlgvqqpaly
```

-continued

```
ABS19067      ~~~~~~~~~mi klqpntviry aldllnevgv ealttrklak rlgvqqpaly
NP 387462     ~~~~~~~~~mn klqreavirt alellndvgm eglttrrlae rlgvqqpaly
NP 387455     ~~~~~~~~~mk klqreavirt alellndvgm eglttrrlae rlgvqqpaly
AAR96033      ~~~~~~~~~mn klqreavirt alellndvgm eglttrrlae rlgvqqpaly
NP 511232     ~~~~~~~~~mn klqreavirt alellndvgm eglttrrlae rlgvqqpaly
AAW83818      ~~~~~~~~~mt klqpntvira aldllnevgv dglttrklae rlgvqqpaly
AAD25094      ~~~~~~~~~mt kldkgtviaa alellnevgm dslttrklae rlkvqqpaly
AB014708      ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~ly
P51560        ~~~~~~~~~mt kldkgtviaa glellnevgm dslttrklae rlkvqqpaly
AAD25537      ~~~~~~~~~mt kldkgtviaa alellnevgm dslttrklae rlrvqqpaly
YP_001220607  ~~~~~~~~~mt kldretviqa alellnevgv dnittrklae rlkvqqpaly
YP_001370475  ~~~~mvsala klhrdaviqt alellnevge eglttrrlae rlgvqqpaly
P21337        ~~~~~~~~~ma rlslddvism altlldsegl eglttrklaq slkieqptly
AAA98409      ~~~~~~~~~ma rlslddvism altlldsegl eglttrklaq slkieqptly
CAC81917      ~~~~~~~~~ma rlslddvism altlldregl eglttrnvaq slkieqptly
P51561        ~~~~~~~~~ma kldkeqvidd alillnevgi eglttrnvaq kigveqptly
ZP_00132379   ~~~~~~~~~ma kldkeqvidn alillnevgi eglttrklaq kigveqptly
AAD12754      ~~~~~~~~~ma kldkeqvidn alillnevgm eglttrklaq klgveqptly
P04483        ~~~~~~~~~ms rldkskvins alellnevgi eglttrklaq klgveqptly
A26948        ~~~~~~~~~ms rldkskvins alellnevgi eglttrklaq klgveqptly
CAC80726      ~~~~~~~~mma rldkkrvies alalldevgm eglttrklaq klnieqpsly
P0ACT4        ~~~~~~~~~MA RLNRESVIDA ALELLNETGI DGLTTRKLAQ KLGIEQPTLY
ZP_01567051   ~~~~~~~~~ma kirrdeivda alalldeggl dalttrrlaq rlgvesaaly
NP_824556     ~~~mvtqrsp kldkkqvvet alrllneagl dgltlraiak elnvqapaly

              51                                                  100
ZP_01558383   whfrnkaell damaeaimle rhgaslprpg dawdawllen arsfrralla
YP_772551     whfrnkaell damaeaimle rhgaslprpg dtwdawllen argfrralla
YP_620166     whfrnkaell damaeaimle rhgeslprpg dvwdvwlaen arsfrralla
EAY62734      whfrnkaell damaeaimle rhgeslprpg dvwdvwlaen arsfrralla
YP_368094     whfrnkgell damaeaimle rhdaslprpg eawdawlldn arsfrralla
AAP93923      whfknkaell kamaetimld .hreevpadm p.wqawvtan ainfrralla
AAW66496      whfknksllm eamaetiine hhlvslpidg mtwqdwllan sysfrralla
CAA24908      whfrnkrall dalaeamlae nhstsvprad ddwrsfltgn arsfrqalla
P03038        whfrnkrall dalaeamlae nhthsvprad ddwrsflign arsfrqalla
ABS19067      whfrnkrall dalaeamlae nhthsvprvd ddwrsflign arsfrqalla
NP_387462     whfknkrall dalaeamlti nhthstprdd ddwrsflkgn acsfrralla
NP_387455     whfrnkrall dalaeamlti nhthstprde ddwrsflkgn acsfrralla
AAR96033      whfknkrall dalaeamlti nhthstprdd ddwrsflkgn acsfrralla
NP_511232     whfknkrall dalaeamlti nhthstprdd ddwrsflkgn acsfrralla
AAW83818      whfrnkrall dalaeamlae nhthsvprad ddwrsflkgn acsfrralla
```

-continued

AAD25094	whfqnkrall	dalaeamlae	rhtrslpeen	edwrvflken	alsfrtalls
AB014708	whfqnkrall	dalaeamlae	rhtrslpeen	edwrvflken	alsfrtalls
P51560	whfqnkrall	dalpeamlre	rhtrslpeen	edwrvflken	alsfrtalls
AAD25537	whfpskrall	dalaeamlte	rhtrslpeen	edwrvflken	alsfrkalls
YP_001220607	whfrnkrall	dalseamlek	nhtrtvpqtg	edwrvflken	alsfrsalls
YP_001370475	whfknkrvll	dalaetilae	hhdhalprag	enwrhflien	ahsfrrallt
P21337	whvrnkqtlm	nmlseailak	hhtrsaplpt	eswqqflgen	alsfrkallv
AAA98409	whlrnkqtlm	nmlseailak	hhtrsaplpt	eswqqflgen	alsfrkallv
CAC81917	whvrnkqtlm	nmlseailak	hhtrsvplpt	eswqqflgen	alsfrkallv
P51561	whvknkrall	dalaetilqk	hhhhvlplpn	etwqdflrnn	aksfrqallm
ZP_00132379	whvknkrall	dalaetilqk	hhhhvlplpn	etwqdflrnn	aksfrqallm
AAD12754	whvknkrall	dalaetilqk	hhhhvlplan	eswqdflrnn	aksfrqallm
P04483	whvknkrall	dalaiemldr	hhthfcpleg	eswqdflrnn	aksfrcalls
A26948	whvknkrall	dalaiemldr	hhthfcpleg	eswqdflrnn	aksfrcalls
CAC80726	whvknkrall	dalsveillr	hhdhfqpqkg	eywadflren	aksfrralls
P0ACT4	WHVKNKRALL	DALAVEILAR	HHDYSLPAAG	ESWQSFLRNN	AMSFRRALLR
ZP_01567051	whyrdksvll	aemaavalar	hhtldvpadt	aqwdawfadn	arsfrralla
NP_824556	whfknkqall	dematemyrr	mtegahlapg	aswgerllhg	nralrtallg

	101				150
ZP_01558383	yrdgarlhag	tr.prtlhfg	sierkvalla	eagfapdeav	dvmyalgrfv
YP_772551	yrdgarlhag	tr.prtlhfd	sierkvalla	dagfapdeav	dvmyalgrfv
YP_620166	yrdgarlhag	tr.pralhfs	sierkvallg	eagfkpdeav	dvmvaigrfv
EAY62734	yrdgarlhag	tr.pralhfs	sierkvallg	eagfkpdeav	dvmvaigrfv
YP_368094	yrdgarlhag	tr.pralhfs	sierkvallg	dagfapdeav	dvmyalgrfv
AAP93923	yrdgarlhag	tr.pgepgfa	iieakvallc	ragftpehav	nllfavgrfv
AAW66496	yrdgarlhar	ts.psqghfn	tieaqvalls	hagfspveav	allmtlgrfi
CAA24908	yrdgarihag	tr.pgapqme	tadaqlrflc	eagfsagdav	nalmtisyft
P03038	yrdgarihag	tr.pgapqme	tadaqlrflc	eagfsagdav	nalmtisyft
ABS19067	yrdgarihag	tr.pgapqme	vvdaqlrflc	eagfsawdav	nalmtisyft
NP_387462	yrdgarihag	tr.saapqme	kadaqlrflc	dagfsagdat	yalmaisyft
NP_387455	yrdgarihag	tr.paapqme	kadaqlrflc	dagflagdat	yalmaisyft
AAR96033	yrdgarihag	tr.paapqme	kadaqlrflc	dagfsagdat	yalmaisyft
NP_511232	yrdgarihag	tr.paapqme	kadaqlrflc	dagfsagdat	yalmaisyft
AAW83818	yrdgarihag	tr.paapqme	kadaqlrflc	dagfsagdat	yalmaisyft
AAD25094	yrdgarihag	tr.ptepnfg	taetqirflc	aegfcpkrav	walrayshyv
AB014708	yrdgarihag	tr.ptepnfg	taetqirflc	aegfcpkrav	walrayshyv
P51560	yrdgarihag	tr.ptepnfg	taetqirflc	aegfcpkrav	walrayshyv
AAD25537	yrdgarihag	tr.ptephyg	taeaqirflc	tagfspkrav	walwayshyv
YP_001220607	yrdgarihag	tr.ptsagye	rvekqirflc	esgfeqpdav	ralvivshyt
YP_001370475	yrdgahihag	tr.pnnnqag	qaetqiefli	qagftpanaa	raliaishyv
P21337	hrdgarlhig	ts.ptppqfe	qaeaqlrcic	dagfsveeal	filqsishft

-continued

```
AAA98409      hrdgarlhig ts.ptppqfe qaeaqlrcic dagfsveeal filqsishft
CAC81917      hrdgarlhig ts.ptppqfe qaeaqlrcic dagfsveeal filqsishft
P51561        yrdggkihag tr.psesqfe tseqqlqflc dagfslsciavyalssiahft
ZP_00132379   yrdggkihag tr.psesqfe tseqqlqflc dagfslsciavyalssiahft
AAD12754      yrdggkihag tr.psancifetseqqlqflc dagftltqav yalssiahft
P04483        hrdgakvhlg tr.ptekqye tlencilaflcqqgfslenal yalsavghft
A26948        hrdgakvhlg tr.ptekqye tlenqlafya nkvfh~~~~~ ~~~~~~~~~~
CAC80726      hrdaakihlg tr.pspecifetveacilaflceqgfsleeal ytlgvvshft
P0ACT4        YRDGAKVHLG TR.PDEKQYD TVETQLRFMT ENGFSLRDGL YAISAVSHFT
ZP_01567051   hrdgarlhag st.pdldave rirpkiaylv raglteqeag mamlaagqft
NP_824556     yrdgakvfsg srftgtehav qleaslrslv eagfdlpqav ratstayfft

              151                                                 200
ZP_01558383   vgwvleeqae aeretd.... .ttlpdtaeh p..llaqgwa alrerggdea
YP_772551     vgwvleeqae aeretd.... .ttlpdtaeh p..llaqgwt alrerggdea
YP_620166     vgwvleeqar pdgdtd.... .allpdaaey p..lfaqgwa alrersgdea
EAY62734      vgwvleeqar pdgdad.... .allpdaaey p..lfakgwa alrersgdea
YP_368094     vgwvleeqae ssdeaa.... .aplpdaaey p..llakgwa alrgrsgdda
AAP93923      vgwvleeqqm qpdda..... .lneadrrry p..llcggwe klqdkgadal
AAW66496      vgwvleeqqe eirsdp.... .pfeadptiy p..lmlqgvn tlqnmnaddi
CAA24908      vgavleeqag dsesgergg. ..tveqapls p..llraaid afdeagpdaa
P03038        vgavleeqag dsdagergg. ..tveqapls p..llraaid afdeagpdaa
ABS19067      vgavleeqag dsdagergg. ..tieqa... p..llravid tfdeagpdav
NP_387462     vgavleqqas eadaeerged qlttsastmp a..rlqsamk ivyeggpdaa
NP_387455     vgavleqqas eadaeerged qlttsastmp a..rlqsamk ivyeggpdaa
AAR96033      vgavleqqas eadaeerged qlttsastmp a..rlqsamk ivyeggpdaa
NP_511232     vgavleqqas eadaeerged qlttsastmp a..rlqsamk ivyeggpdaa
AAW83818      vgavleqqas eadaeerged qlttsastmp a..rlqsamk ivyeggpdaa
AAD25094      vgsvleqqas ..dadervpd rpdvseqaps s..flhdlfh eletdgmdaa
AB014708      vgsvleqqas ..dadervpd rpdvseqaps s..flhdlfh eletdgmdaa
P51560        vgsvleqqas ..dadervpd rpdvseqaps s..flhvlfh eletdgmdaa
AAD25537      vgsvleqqas ..nandrmsd ksdvseqaps s..flhdlfh eletdgmdap
YP_001220607  tgsyseqqaa ledsserkqa skeapaq.ps q..flshafd tfdaegadfa
YP_001370475  vgsaleqqa. ..dihesvpg ..daysitat s..eiaqaia ildadgaenl
P21337        lgavleeq.. atnqien..n hvid...aap p..llqeafn iqartsaema
AAA98409      lgavleeq.. atnqien..n hvid...aap p..llqeafn iqartsaema
CAC81917      lgavleeq.. atnptey..n tvmd...avp p..llqeafn vqtrttaeta
P51561        lgsvletqeh qesqker..e kvetdtvayp p..lltqava imdsdngdaa
ZP_00132379   lgsvletqeh qesqker..e kvetdtvayp p..lltqava imdsdngdaa
AAD12754      lgsvletqeh qesqker..e kvpkteinyp p..lltqaid imdsdngeaa
P04483        lgcvledgeh qvakeer..e tpttdsm..p p..llrqaie lfdhqgaepa
```

-continued

```
A26948          ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~

CAC80726        lgsvleerey leamrdd..d paihaam..p p..lltkale imeqdtgekp

POACT4          LGAVLEQQEH TAALTDR..P AAPDENL..P P..LLREALQ IMDSDDGEQA

ZP_01567051     igcvleqqaa qgrgaeepar rdaaddrp.. .....rtsga alapidpgva

NP_824556       lgfvteeqgv eplpgerreg ydvderaarm adfplaaaag aeifqnyeeg

                201                                   237
ZP_01558383     fergvalivd gararla.ar rrgg~~~~~~ ~~~~~~~

YP_772551       fergvalivd gararla.ar qrgg~~~~~~ ~~~~~~~

YP_620166       fergiawivd gararla.ar rag~~~~~~~ ~~~~~~~

EAY62734        fergiawivd gararla.ar rag~~~~~~~ ~~~~~~~

YP_368094       fergvawivd gararla.ar erg~~~~~~~ ~~~~~~~

AAP93923        feaglrllvd gaeaaltnan nhgaqs~~~~ ~~~~~~~

AAW66496        fengirmvii gaerqldikm qt~~~~~~~~ ~~~~~~~

CAA24908        feqglavivd glakrrlvvr nvegprkgdd ~~~~~~~

P03038          feqglavivd glakrrlvvr nvegprkgdd ~~~~~~~

ABS19067        felglavivd glakrrlvar niqgprkgdd ~~~~~~~

NP_387462       ferglaliig gleqvrlspa sspagrtnlv lalaags

NP_387455       ferglaliig gleqvrlspa sspagrtnlv lalaags

AAR96033        ferglaliig glersacais ll~~~~~~~~ ~~~~~~~

NP_511232       ferglaliig glekmrlttn dievlknvde ~~~~~~~

AAW83818        ferglaliig glersacais ll~~~~~~~~ ~~~~~~~

AAD25094        fnfgldslia gferlrss.. ttd~~~~~~~ ~~~~~~~

AB014708        fnfgldslia gferlrss.. ttd~~~~~~~ ~~~~~~~

P51560          fnfgldslia gferlraavl atd~~~~~~~ ~~~~~~~

AAD25537        fnfgldslia gfeqlrls.. ttd~~~~~~~ ~~~~~~~

YP_001220607    feygldalis glemkkatk~ ~~~~~~~~~~ ~~~~~~~

YP_001370475    fdfglmllvd glerhrqs~~ ~~~~~~~~~~ ~~~~~~~

P21337          fhfglkslif gfsaqldekk htpiedgnk~ ~~~~~~~

AAA98409        fhfglkslif gfsaqldekk htpiedgnk~ ~~~~~~~

CAC81917        fhfglrsliv gfsaqlde.k ymsiqgnnk~ ~~~~~~~

P51561          flfvldvmis gletvlksak ~~~~~~~~~~ ~~~~~~~

ZP_00132379     flfvldvmis gletvlksak ~~~~~~~~~~ ~~~~~~~

AAD12754        flfvldvmis gletvinnhh ~~~~~~~~~~ ~~~~~~~

P04483          flfgleliic glekqlkces gs~~~~~~~~ ~~~~~~~

A26948          ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~~~~ ~~~~~~~

CAC80726        flfgleviil gleakqkqkk gnqe~~~~~~ ~~~~~~~

P0ACT4          FLHGLESLIR GFEVQLTALL QIVGGDKLII PFC~~~~

ZP_01567051     fefglglivd glrrrvdra~ ~~~~~~~~~~ ~~~~~~~

NP_824556       feeglrlvia giearygir~ ~~~~~~~~~~ ~~~~~~~
```

Amino acid positions having relatively conserved amino acid substitutions between family members were considered for harvesting diversity. In addition, positions were chosen for variation based on spacing to limit the number of changes in a pair of overlapping oligonucleotides. A summary of the library is shown in Table 11. The objective of this library was to recover hits improved for reactivity to either ethametsulfuron or chlorsulfuron.

TABLE 11

	Residue Position																			
	52	55	62	69	73	76	79	83	85	88	93	96	98	101	102	109	110	114	117	120	125
L1-9 Backbone Sequence	L	L	R	P	E	Q	L	A	S	C	H	G	K	L	G	Q	Y	E	L	L	F
Shuffling Diversity	L M	L M	D E K R	L P	A D E	A D Q R	F L V	A N S	A G N S	C R S	H Y	A G	K R	I L S	G R	H N Q	F V Y	D E R	I L V	L M Y	F L

	Residue Position															
	129	130	137	140	145	149	162	167	170	175	181	183	189	190	193	197
L1-9 Backbone Sequence	N	A	V	F	V	Q	T	P	L	E	G	E	G	L	I	L
Shuffling Diversity	D E H N Q	A G	A I L V	F Y	A V	Q R	Q T	P S	F I L M	D E N	G S	D E G	G V	I L V	I L M V	A F I L

Assembly of the L4 library synthetic oligonucleotides was done as for the previous libraries, except that two sets of oligonucleotide pools were used. First, multiple oligonucleotides representing diversity at a single oligonucleotide annealing location are pooled ("Group" in Table 12). Next, an equal volume of each group of oligos is pooled to represent the novel L4 diversity. Likewise, oligonucleotides representing the L1-9 backbone sequence were pooled (Table 13). The L4 assembly reaction was carried out by spiking the oligonucleotide diversity pool into the L1-9 backbone pool at an approximately 1:3 ratio.

TABLE 12

Oligo	Sequence	Group	SEQ ID
L4:01	TATTGGCATGTAAAAAATAAGCGAGCTCTGWTGGACGCCWTG	1	987
L4:02	TATTGGCATGTAAAGAATAAGVGCGCTCTGWTGGACGCCWTG	1	988
L4:03	GCCATTGAGATGCTCGATARACACGCCACTCACTTTTGCCYC	2	989
L4:04	GCCATTGAGATGCTCGATGAKCACGCCACTCACTTTTGCCYC	2	990
L4:05	TTAGAAGGGGMWAGCTGGCAAGATTTTBTTCGTAATAACGCA	3	991
L4:06	TTAGAAGGGGMWAGCTGGCAAGATTTTBTTCGTAATAACART	3	992
L4:07	TTAGAAGGGGMWAGCTGGAGGGATTTTBTTCGTAATAACGCA	3	993
L4:08	TTAGAAGGGGMWAGCTGGAGGGATTTTBTTCGTAATAACART	3	994
L4:09	TTAGAAGGGGMWAGCTGGGMTGATTTTBTTCGTAATAACGCA	3	995
L4:10	TTAGAAGGGGMWAGCTGGGMTGATTTTBTTCGTAATAACART	3	996
L4:11	AAAARTATGAGAHGTGCTTTACTAAGTYACCGCGATGSAGCA	4	997
L4:12	AAAGSAATGAGAHGTGCTTTACTAAGTYACCGCGATGSAGCA	4	998
L4:13	ARAGTATGCTCCRGGACAGGATTTACAGAAAAACAAKTTGAA	5	999
L4:14	ARAGTATGCTCCRGGACAGGATTTACAGAAAAACAATACGAA	5	1000
L4:15	ARAGTATGCTCCRGGACAGGATTTACAGAAAAAMATKTTGAA	5	1001
L4:16	ARAGTATGCTCCRGGACAGGATTTACAGAAAAAMATTACGAA	5	1002
L4:17	ARAGTATGCMTCRGGACAGGATTTACAGAAAAACAAKTTGAA	5	1003
L4:18	ARAGTATGCMTCRGGACAGGATTTACAGAAAAACAATACGAA	5	1004
L4:19	ARAGTATGCMTCRGGACAGGATTTACAGAAAAAMATKTTGAA	5	1005
L4:20	ARAGTATGCMTCRGGACAGGATTTACAGAAAAAMATTACGAA	5	1006
L4:21	ACTGCTGAMAATTCAVTTGCCTTTMTGTGCCAACAAGGTTTK	6	1007
L4:22	ACTGCTGAMAATTCAVTTGCCTTTTACTGCCAACAAGGTTTK	6	1008

TABLE 12-continued

Oligo	Sequence	Group	SEQ ID
L4:23	ACTGCTAGGAATTCAVTTGCCTTTMTGTGCCAACA AGGTTTK	6	1009
L4:24	ACTGCTAGGAATTCAVTTGCCTTTTACTGCCAACA AGGTTTK	6	1010
L4:25	TCACTAGAGVACGSATTATATGCAATGCAAGCTGC ATGTATT	7	1011
L4:26	TCACTAGAGVACGSATTATATGCAATGCAAGCTVT CTGTATT	7	1012
L4:27	TCACTAGAGSAAGSATTATATGCAATGCAAGCTGC ATGTATT	7	1013
L4:28	TCACTAGAGSAAGSATTATATGCAATGCAAGCTVT CTGTATT	7	1014
L4:29	TWCACTTTAGGTTGCGYATTGCTCGATCAAGAGTT GCAAGTC	8	1015
L4:30	TWCACTTTAGGTTGCGYATTGCTCGATCGTGAGTT GCAAGTC	8	1016
L4:31	GCTAAAGAAGAAAGGGAAACACCTCAAACTGATAG TATGYCT	9	1017
L4:32	GCTAAAGAAGAAAGGGAAACACCTACTACTGATAG TATGYCT	9	1018
L4:33	CCATTAWTKCGACAAGCTTTGAATTTAAAGGATCA CCAARGC	10	1019
L4:34	CCATTAWTKCGACAAGCTTTGGAWTTAAAGGATCA CCAARGC	10	1020
L4:35	GCAGRWCCAGCCTTCTTATTCGKGVTTGAATTGVT KATATGC	11	1021
L4:36	GGAHTTGAAAAACAACTTAAATGTGAAAGTGGGTC TTAA	12	1022
L4:37	GGAGCTGAAAAACAACTTAAATGTGAAAGTGGGTC TTAA	12	1023
L4:38	TYTATCGAGCATCTCAATGGCCAWGGCGTCCAWCA GAGCTCG	13	1024
L4:39	MTCATCGAGCATCTCAATGGCCAWGGCGTCCAWCA GAGCTCG	13	1025
L4:40	TYTATCGAGCATCTCAATGGCCAWGGCGTCCAWCA GAGCGCB	13	1026
L4:41	MTCATCGAGCATCTCAATGGCCAWGGCGTCCAWCA GAGCGCB	13	1027
L4:42	TTGCCAGCTWKCCCCTTCTAAGRGGCAAAAGTGAG TGGCGTG	14	1028
L4:43	CCTCCAGCTWKCCCCTTCTAAGRGGCAAAAGTGAG TGGCGTG	14	1029
L4:44	AKCCCAGCTWKCCCCTTCTAAGRGGCAAAAGTGAG TGGCGTG	14	1030
L4:45	TAAAGCACDTCTCATAYTTTTTGCGTTATTACGAA VAAAATC	15	1031
L4:46	TAAAGCACDTCTCATTSCTTTTGCGTTATTACGAA VAAAATC	15	1032
L4:47	TAAAGCACDTCTCATAYTTTTAYTGTTATTACGAA VAAAATC	15	1033
L4:48	TAAAGCACDTCTCATTSCTTTAYTGTTATTACGAA VAAAATC	15	1034

TABLE 12-continued

Oligo	Sequence	Group	SEQ ID
L4:49	TCCTGTCCYGGAGCATACTYTTGCTSCATCGCGGT RACTTAG	16	1035
L4:50	TCCTGTCCYGAKGCATACTYTTGCTSCATCGCGGT RACTTAG	16	1036
L4:51	GGCAABTGAATTKTCAGCAGTTTCAAMTTGTTTTT CTGTAAA	17	1037
L4:52	GGCAABTGAATTCCTAGCAGTTTCAAMTTGTTTTT CTGTAAA	17	1038
L4:53	GGCAABTGAATTKTCAGCAGTTTCGTATTGTTTTT CTGTAAA	17	1039
L4:54	GGCAABTGAATTCCTAGCAGTTTCGTATTGTTTTT CTGTAAA	17	1040
L4:55	GGCAABTGAATTKTCAGCAGTTTCAAMATKTTTTT CTGTAAA	17	1041
L4:56	GGCAABTGAATTCCTAGCAGTTTCAAMATKTTTTT CTGTAAA	17	1042
L4:57	GGCAABTGAATTKTCAGCAGTTTCGTAATKTTTTT CTGTAAA	17	1043
L4:58	GGCAABTGAATTCCTAGCAGTTTCGTAATKTTTTT CTGTAAA	17	1044
L4:59	ATATAATSCGTBCTCTAGTGAMAAACCTTGTTGGC ACAKAAA	18	1045
L4:60	ATATAATSCTTSCTCTAGTGAMAAACCTTGTTGGC ACAKAAA	18	1046
L4:61	ATATAATSCGTBCTCTAGTGAMAAACCTTGTTGGC AGTAAAA	18	1047
L4:62	ATATAATSCTTSCTCTAGTGAMAAACCTTGTTGGC AGTAAAA	18	1048
L4:63	CAATRCGCAACCTAAAGTGWAAATACATGCAGCTT GCATTGC	19	1049
L4:64	CAATRCGCAACCTAAAGTGWAAATACAGABAGCTT GCATTGC	19	1050
L4:65	TGTTTCCCTTTCTTCTTTAGCGACTTGCAACTCTT GATCGAG	20	1051
L4:66	TGTTTCCCTTTCTTCTTTAGCGACTTGCAACTCAC GATCGAG	20	1052
L4:67	CAAAGCTTGTCGMAWTAATGGAGRCATACTATCAG TTTGAGG	21	1053
L4:68	CAAAGCTTGTCGMAWTAATGGAGRCATACTATCAG TAGTAGG	21	1054
L4:69	GAATAAGAAGGCTGGWYCTGCGCYTTGGTGATCCT TTAAATT	22	1055
L4:70	GAATAAGAAGGCTGGWYCTGCGCYTTGGTGATCCT TTAAWTC	22	1056
L4:71	TTTAAGTTGTTTTTCAADTCCGCATATMABCAATT CAABCMC	23	1057
L4:72	TTTAAGTTGTTTTTCAGCTCCGCATATMABCAATT CAABCMC	23	1058
L1:50	GGGAACTTCGGCGCGCCTTAAGACCCACTTTCACA	24	1059

TABLE 13

Oligo	Sequence	SEQ ID
L1-9:01	TATTGGCATGTAAAGAATAAGCGTGCTCTGTTGGAC GCCCTG	1060
L1-9:02	GCCATTGAGATGCTCGATCGTCACGCCACTCACTTT TGCCCT	1061
L1-9:03	TTAGAAGGGGAAAGCTGGCAAGATTTTCTCCGTAAT AATGCA	1062
L1-9:04	AAATCAATGAGATGCGCTTTACTAAGTCATCGCGAT GGGGCA	1063
L1-9:05	AAGGTATGTCTTGGTACAGGATTCACAGAAAAACAG TACGAA	1064
L1-9:06	ACTGCTGAAAATAGTTTGGCCTTTCTGTGCCAACAA GGTTTC	1065
L1-9:07	TCACTAGAGAATGCTTTATATGCAATGCAAGCTGTC TGTATC	1066
L1-9:08	TTCACTTTAGGTTGCGTTTTGCTGGATCAAGAGCTC CAAGTC	1067
L1-9:09	GCTAAAGAAGAAAGGGAAACACCTACTACTGATAGT ATGCCC	1068
L1-9:10	CCATTATTGCGACAAGCTTTGGAATTAAAAGATCAC CAAGGG	1069
L1-9:11	GCAGAGCCAGCCTTCTTATTCGGATTGGAATTGATA ATATGC	1070
L1-9:12	GGATTGGAAAAACAACTTAAATGTGAAAGTGGGTCT TAA	1071
L1-9:13	ACGATCGAGCATCTCAATGGCCAGGGCGTCCAACAG AGCACG	1072
L1-9:14	TTGCCAGCTTTCCCCTTCTAAAGGGCAAAAGTGAGT GGCGTG	1073
L1-9:15	TAAAGCGCATCTCATTGATTTTGCATTATTACGGAG AAAATC	1074
L1-9:16	TCCTGTACCAAGACATACCTTTGCCCCATCGCGATG ACTTAG	1075
L1-9:17	GGCCAAACTATTTTCAGCAGTTTCGTACTGTTTTTC TGTGAA	1076
L1-9:18	ATATAAAGCATTCTCTAGTGAGAAACCTTGTTGGCA CAGAAA	1077
L1-9:19	CAAAACGCAACCTAAAGTGAAGATACAGACAGCTTG CATTGC	1078
L1-9:20	TGTTTCCCTTTCTTCTTTAGCGACTTGGAGCTCTTG ATCCAG	1079
L1-9:21	CAAAGCTTGTCGCAATAATGGGGGCATACTATCAGT AGTAGG	1080
L1-9:22	GAATAAGAAGGCTGGCTCTGCCCCTTGGTGATCTTT TAATTC	1081
L1-9:23	TTTAAGTTGTTTTTCCAATCCGCATATTATCAATTC CAATCC	1082
L1-9:24	GGGAACTTCGGCGCGCCTTAAGACCCACTTTCACA	1083

The assembly reaction products were cloned into the pVER7314 backbone and transformed into tester strain *E. coli* KM3. To carry out library diversity analysis, DNA preps from 96 colonies grown on LB+Cb only (representing no repressor positive selection bias) were subjected to sequence analysis. These data showed approximately 30% of the clones recovered were unaltered L1-9 backbone and the remaining clones had approximately one to two targeted changes per clone. Additional non-targeted residue changes were recovered in the mutated population, either due to PCR errors or from poor quality oligonucleotides incorporated into the assembly reactions.

I. Library L4 Screening

Approximately 20,000 clones arising from the repressor prescreen were tested for activation by 0, 0.2 and 1 μg/ml concentrations of ethametsulfuron using the M9 assay plates. Surprisingly, over 100 hits were observed from the 0.2 μg/ml ethametsulfuron treatment. These putative hits were re-arrayed in 96-well format and re-tested for β-galactosidase induction by 0, 0.2 and 1 ppm ethametsulfuron using a liquid culture based assay system. FIG. **4** shows relative β-galactosidase activities of 45 exemplary putative library L4 hit clones 97-142 against 0, 0.2 and 1 ppm ethametsulfuron. Cultures grown in 96-well format were subcultured into fresh LB with inducer at the indicated concentration and grown overnight and then processed for the enzyme assay. For detection of induced activity, 5 μl of perforated cell mixture was used. For detection of background activity, 25 μl of cells were used such that detectable activity could be measured in the same time frame for all treatments. Background activity values were multiplied by ten to bring them into the range of the graph. The numbers below each sample refers to the library clone number. The latter part of the graph contains the controls 1st round hit L1-9 as well as wt TetR.

DNA sequences for all 142 putative hit clones were determined and the translated polypeptides aligned. After assigning each polypeptide in the alignment with a relative ethametsulfuron response, patterns of amino acid incorporation at varied or mutated residues associated with high or low response activity and high or low uninduced activity were identified. The most significant findings from this analysis were: C138G or L170V mutations were heavily favored in the top clones L4-59, -89, -110, -116, -118, -120, -124, -129, -133, -139, -140 and -142; and K108Q was heavily incorporated in clones with the highest activity at the lowest dose of 0.2 ppm, but these clones showed somewhat leaky background (e.g., L4-98, -106, -113, -126, -130, and -141). Results from clone L4-18 having the K108Q shows another possible interesting mutation of L55M. This clone is induced to a high level with 0.2 ppm Es, but does not show the associated high background activity typically observed for K108Q-containing clones. The L55M mutation may have increased repressor activity. It is of interest that none of these changes other than L55M were designed diversity—all were derived from false incorporation of nucleotides during library assembly and few of these changes were represented in the unselected clone population.

J. Third Round Library Designs and Screening

Library L6: Shuffling for Enhanced Chlorsulfuron Response

Since clones L2-14 and L2-18 had the best chlorsulfuron activity profile from library L2, their amino acid diversity was used as the basis for the next round of shuffling. In addition to the diversity provided by these backbone sequences, additional residue changes thought to enhance packing of chlorsulfuron based on the 3D model predictions were included. New amino acid positions targeted were 67, 109, 112 and 173 (see, Table 14). Substitution of Gln (Q) at position 108 and Val (V) at position 170 were shown to likely be important changes in library L4 for gaining enhanced SU responsiveness and so were varied here as well. A summary of the diversity chose is shown in Table 14. The oligonucleotides designed and used to generate library 6 are shown in Table 15.

Library L6 was assembled, rescued, ligated into pVER7314, transformed into *E. coli* KM3 and plated out onto LB carbenicillin/kanamycin, and carbenicillin only control media as before. Library plates were then picked into 42 384-well microtiter plates (~16,000 clones) containing 60 µl LB carbenicillin (Cb) broth per well. After overnight growth at 37° C. the cultures were stamped onto M9 assay plates containing no inducer, 0.2 ppm, and 2.0 ppm chlorsulfuron as test inducer. Following incubation at 30° C. for ~48 hrs, putative hits responding to chlorsulfuron treatment as determined by increased blue colony color were re-arrayed into six 96-well microtiter plates and used to stamp a fresh set of M9 assay plates to confirm the above results. For a more detailed analysis of the relative induction by chlorsulfuron, digital photographs were taken of the plates after various time points of incubation at 30° C. and colony color intensity measured using the digital image analysis freeware program ImageJ (Rasband, US National Institutes of Health, Bethesda, Md., USA, rsb.info.nih.gov/ij/, 1997-2007). Using these results enabled ranking of clones in multiplex format by background activity (no inducer), activation with low or high level inducer application (blue color with inducer), and fold activation (activation divided by background). Activation studies using 0.2 µg/ml chlorsulfuron as inducer for the top set of clones shows an approximately 3 fold improvement in activation while obtaining lower un-induced levels of expression (Table 12). In addition to this analysis, DNA sequence information for most clones (490 clones) was obtained and the deduced polypeptides aligned with each other as well as with their corresponding activity information. From this analysis sequence-activity relationships were derived (Table 12). Residues biased for improved activity are indicated in larger bold type. Briefly, C at position 100, and Q at positions 108 and 109 strongly correlated with activation, while R at position 138, L at position 170, and A or G at position 173 were highly preferred in clones with the lowest background activity. Though some positions were strongly biased, i.e., observed more frequently in the selected population, the entirety of introduced diversity was observed in the full hit population. This information will aid in the design of further libraries to improve responsiveness to chlorsulfuron.

TABLE 14

Sequence	Amino acid residue position														
Name	60	64	67	82	86	100	105	108	109	112	113	116	134	138	139
Library Diversity		A Q	M Y F L I V	N T		C W		Q K	M L Q H	S T A G		M Q	M V	G R	N V
wt reference	L	H	F	N	F	H	P	K	Q	T	L	Q	L	G	H
L2-14	M	A	F	N	M	C	W	K	Q	T	A	M	V	R	V
L2-18	M	Q	F	T	M	W	W	K	Q	T	A	Q	M	R	N
L6-1B03	M	A	I	N	M	C	W	Q	Q	A	A	M	V	R	V
L6-2C09	M	Q	Y	T	M	C	W	Q	L	T	A	Q	M	R	V
L6-2D07	M	Q	F	T	M	C	W	Q	Q	T	A	M	M	R	V
L6-3H02	M	A	Y	T	M	C	W	Q	H	S	A	M	V	R	V
L6-4D10	M	Q	Y	N	M	C	W	K	Q	S	A	M	V	R	V
L6-5F05	M	A	I	N	M	C	W	Q	Q	A	A	Q	V	R	V
L6-5G06	M	Q	Y	N	M	C	W	Q	Q	T	A	Q	V	R	V
L6-5H06	M	Q	I	N	M	C	W	K	Q	T	A	M	V	R	V
L6-5H12	M	A	Y	N	M	C	W	K	Q	T	A	Q	M	R	V
L6-6F07	M	A	L	T	M	C	W	Q	Q	S	A	M	M	R	V
Bias in top population		none	Y	N		**C**		Q	**Q**	none		none	V	**R**	V

Sequence	Amino acid residue position									
Name	147	151	164	170	173	174	177	178		
Library Diversity		S L		L A V	G A V	L W			0.2 ppm 48 hr	Control 84 hr	0.2 ppm 48 hr/ Control 84 hr
wt reference	E	H	D	L	A	I	F	D	5.2	5.3	1.0
L2-14	F	S	A	L	A	L	K	D	11.8	6.6	1.8
L2-18	F	L	A	L	A	W	K	D	5.9	5.7	1.0
L6-1B03	F	S	A	L	A	W	K	D	30.0	6.6	4.6
L6-2C09	F	L	A	L	A	W	K	D	13.6	5.2	2.6
L6-2D07	F	S	A	V	A	W	K	D	20.0	5.8	3.4
L6-3H02	F	S	A	V	A	W	K	V	15.8	5.6	2.8
L6-4D10	F	S	A	L	A	W	K	D	18.4	5.0	3.7
L6-5F05	F	L	A	L	A	W	K	D	22.0	5.4	4.1
L6-5G06	F	L	A	L	G	W	K	D	34.4	7.0	4.9
L6-5H06	F	L	A	V	A	W	K	D	13.7	5.1	2.7

TABLE 14-continued

L6-5H12	F	L	A	V	A	W	K	D	23.7	5.7	4.2
L6-6F07	F	S	A	L	A	W	K	D	11.6	5.1	2.3
Bias in top population		none		**L**	**A/G**	W		D			

TABLE 15

Oligo	Sequence	SEQ ID
L6:1	TATTGGCATGTAAAAAATAAGCGAGCTCTGCTCGACGCCTTA	1084
L6:2	GCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATATGC	1085
L6:3	ATATAATGCATTCTCTAGTGAAAAACCTTGTTGGCATAAAAA	1086
L6:4	TTTAAGTTGTTTTTCTAATCCGCATATGATCAATTCAAGGCC	1087
L6:5	TTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAMTGCT	1088
L6:6	TAAAGCACATCTCATACTTTTAGCAKTATTACGTAAAAAATC	1089
L6:7	TTGCCAGCTTTCCCCTTCTAAAGGGCAMAHGTGAGTTGCGTG	1090
L6:8	TTGCCAGCTTTCCCCTTCTAAAGGGCAATAGTGAGTTGCGTG	1091
L6:9	GAATAAGAAGGCTGGCTCTGCACCTTGGTGATCCTTTAATTC	1092
L6:10	GCCATTGAGATGATGGATAGGCACGCAACTCACTATTGCCCT	1093
L6:11	RSTGCTGAAAATATGTTAGCCTTTTTATGCCAACAAGGTTTT	1094
L6:12	TTTACTTTAGGTTGCGTATTGTTTGATCAAGAGCTCCAAGTC	1095
L6:13	TGTTTCCCTTTCTTCTTTAGCGACTTGGAGCTCTTGATCAAA	1096
L6:14	GCCATTGAGATGATGGATAGGCACGCAACTCACDTKTGCCCT	1097
L6:15	GCCATTGAGATGATGGATAGGCACCAAACTCACDTKTGCCCT	1098
L6:16	GCCATTGAGATGATGGATAGGCACCAAACTCACTATTGCCCT	1099
L6:17	AAAAGTATGAGATGTGCTTTACTAAGCCATCGCGATGGAGCA	1100
L6:18	AAAGTATGKTTAGGTACACGCTGGACAGAAMAACAWTATGAA	1101
L6:19	AAAGTATGKTTAGGTACACGCTGGACAGAAMAAWTGTATGAA	1102
L6:20	RSTGCTGAAAATCAATTAGCCTTTTTATGCCAACAAGGTTTT	1103
L6:21	TCACTAGAGAATGCATTATATGCARTGAGTGCGTGGRGGGTG	1104
L6:22	TCACTAGAGAATGCATTATATGCARTGAGTGCGTGGRGGAAC	1105

TABLE 15-continued

Oligo	Sequence	SEQ ID
L6:23	TTTACTTTAGGTTGCGTATTGTTTGATCAAGAGAGCCAAGTC	1106
L6:24	GCTAAAGAAGAAAGGGAAACACCTACTACTGCTAGTATGCCG	1107
L6:25	CCATTAKTGCGACAAGBTTKGGAATTAAAGGATCACCAAGGT	1108
L6:26	CCATTAGCCCGACAAGBTTKGGAATTAAAGGATCACCAAGGT	1109
L6:27	GGATTAGAAAAACAACTTAAATGCGAAAGTGGGTCTTAA	1110
L6:28	CCTATCCATCATCTCAATGGCTAAGGCGTCGAGCAGAGCTCG	1111
L6:29	TTGCCAGCTTTCCCCTTCTAAAGGGCAMAHGTGAGTTTGGTG	1112
L6:30	TTGCCAGCTTTCCCCTTCTAAAGGGCAATAGTGAGTTTGGTG	1113
L6:31	GCGTGTACCTAAMCATACTTTTGCTCCATCGCGATGGCTTAG	1114
L6:32	GGCTAACATATTTTCAGCASYTTCATAWTGTTKTTCTGTCCA	1115
L6:33	GGCTAATTGATTTTCAGCASYTTCATAWTGTTKTTCTGTCCA	1116
L6:34	GGCTAACATATTTTCAGCASYTTCATACAWTTKTTCTGTCCA	1117
L6:35	GGCTAATTGATTTTCAGCASYTTCATACAWTTKTTCTGTCCA	1118
L6:36	CAATACGCAACCTAAAGTAAACACCCYCACAGCACTCAYTGC	1119
L6:37	CAATACGCAACCTAAAGTAAAGTTCCYCACAGCACTCAYTGC	1120
L6:38	TGTTTCCCTTTCTTCTTTAGCGACTTGGCTCTCTTGATCAAA	1121
L6:39	CMAAVCTTGTCGCAMTAATGGCGGCATACTAGCAGTAGTAGG	1122
L6:40	CMAAVCTTGTCGGGCTAATGGCGGCATACTAGCAGTAGTAGG	1123
L6:41	GGGAACTTCGGCGCGCCTTAAGACCCACTTTCGCA	1124

K. Library L7: Shuffling for Enhanced Ethametsulfuron Response

The choice of parents to represent the amino acid residue diversity for library L7 was based on the conclusions of library L4 analysis—namely incorporation of mutations K108Q, C138G and L170V. Clones were also chosen to bring in other changes that occurred at a much lower frequency in L4, but may have been contributing to activity. These residues are L55M, N129H, V137A and F140Y. In addition to family diversity, other residue modifications were introduced at

amino acid positions 67, 109, 112, 117, 131 and 173 based on structural modeling. This information is summarized in Table 14 showing L7 diversity summary. Also shown in Table 16 is a sequence alignment the top 10 performing L7 hits limited to the differences between the hits and wt TetR. Activity was determined using image analysis of colony color (ImageJ software) on M9 assay plates containing 0, 0.02 or 0.2 ppm ethametsulfuron. At the bottom of Table 16 is a summary of the sequence-activity relationship analysis for the entire data set derived from more than 300 clones, with the strongly biased positions shown in larger bolded type. Even though some positions were strongly biased, i.e., observed more frequently in the selected population, e.g., M at position 55, the entirety of introduced diversity was observed in the full hit population.

TABLE 16

	Amino acid residue position																										
Sequence	55	64	67	85	86	100	104	105	108	109	112	113	116	117	129	131	134	135	137	138	139	140	147	151	170	173	174	177
wt reference	L	H	F	S	F	H	R	P	K	Q	T	L	Q	L	N	L	L	S	V	G	H	F	E	H	L	A	I	F
L7 diversity	M L		M Y F L V						Q K	M L Q H	S T A G			M L	N H	M L			A V	G C		F Y			L A V	G A V		
L7-1C03-A05	M	A	V	I	M	C	G	F	Q	Q	–	A	S	–	H	–	M	Q	A	G	I	–	L	L	V	–	L	N
L7-1C07-A06	M	A	Y	–	M	C	G	F	Q	Q	–	A	S	–	H	–	M	Q	A	G	I	–	L	L	V	–	L	K
L7-1F08-A11	M	A	Y	–	M	C	G	F	Q	Q	–	A	S	–	–	–	M	Q	–	G	I	Y	L	L	V	–	L	K
L7-1G06-B02	M	A	Y	–	M	C	G	F	Q	Q	–	A	S	M	–	–	M	Q	–	G	I	Y	L	L	V	–	L	K
L7-2C11-B11	M	A	Y	–	M	C	G	F	Q	Q	–	A	S	M	–	M	M	Q	A	G	I	–	L	L	V	V	L	K
L7-2D08-C02	M	A	Y	–	M	C	G	F	Q	Q	–	A	S	–	H	–	M	Q	–	G	I	Y	L	L	V	–	L	K
L7-3A10-C09	M	A	Y	–	M	C	G	F	Q	Q	–	A	S	–	H	–	M	Q	–	G	I	Y	L	L	V	–	L	K
L7-3C08-C10	M	A	Y	–	M	C	G	F	Q	Q	G	A	S	M	–	M	M	Q	A	G	I	Y	L	L	V	–	L	K
L7-3E03-D01	M	A	Y	–	M	C	G	F	Q	Q	–	A	S	–	H	M	M	Q	A	G	I	Y	L	L	V	–	L	K
L7-3E04-D02	M	A	Y	–	M	C	G	F	Q	Q	S	A	S	–	H	M	M	Q	A	G	I	–	L	L	V	–	L	K
Bias in top population	**M**		Y						**Q**	**Q**					N					G					**V**			

The L7 library was constructed as for Library L1 using the set of oligonucleotides shown below in Table 17.

TABLE 17

Oligo ID	Oligo sequence	SEQ ID
L7:01	TATTGGCATGTAAAAAATAAGCGAGCTCTGCTCGACGCA WTG	1125
L7:02	GCCATTGAGATGCTGGATAGGCACGCGACTCACDTSTGC CCT	1126
L7:03	GCCATTGAGATGCTGGATAGGCACGCGACTCACTATTGC CCT	1127
L7:04	TTAGAAGGGGAAAGCTGGCAAGATTTTTTACGTAATAAC GCT	1128
L7:05	AAAAGTATGAGATGTGCTTTACTAAGTCATCGCGATGGA GCA	1129
L7:06	AAAGTATGTTTAGGTACAGGCTTTACAGAAMAGMTGTAT GAA	1130
L7:07	AAAGTATGTTTAGGTACAGGCTTTACAGAAMAGCAMTAT GAA	1131
L7:08	RSTGCCGAAAATAGTMTGGCCTTTTTATGCCAACAAGGT TTT	1132
L7:09	TCACTAGAGMACGCAMTGTATGCAATGCAGGCTGYTKGT ATT	1133
L7:10	TWTACTTTAGGTTGCGTATTGTTGGATCAAGAGCTTCAA GTC	1134
L7:11	GCTAAAGAAGAAAGGGAAACACCTACTACTGATAGTATG CCG	1135

TABLE 17-continued

Oligo ID	Oligo sequence	SEQ ID
L7:12	CCATTAGCTCGACAAGBTCTGGAATTAAAGGATCACCAA GGT	1136
L7:13	CCATTASTCCGACAAGBTCTGGAATTAAAGGATCACCAA GGT	1137
L7:14	GCAGAGCCAGCCTTCTTATTCGGCCTTGAATTGATCATA TGC	1138
L7:15	GGATTAGAAAAACAACTTAAATGTGAAAGTGGGTCTTAA	1139
L7:16	CCTATCCAGCATCTCAATGGCCAWTGCGTCGAGCAGAGC TCG	1140
L7:17	TTGCCAGCTTTCCCCTTCTAAAGGGCASAHGTGAGTCGC GTG	1141
L7:18	TTGCCAGCTTTCCCCTTCTAAAGGGCAATAGTGAGTCGC GTG	1142
L7:19	TAAAGCACATCTCATACTTTTAGCGTTATTACGTAAAAA ATC	1143
L7:20	GCCTGTACCTAAACATACTTTTGCTCCATCGCGATGACT TAG	1144

TABLE 17-continued

Oligo ID	Oligo sequence	SEQ ID
L7:21	GGCCAKACTATTTTCGGCASYTTCATACAKCTKTTCTGT AAA	1145
L7:22	GGCCAKACTATTTTCGGCASYTTCATAKTGCTKTTCTGT AAA	1146
L7:23	ATACAKTGCGTKCTCTAGTGAAAAACCTTGTTGGCATAA AAA	1147
L7:24	CAATACGCAACCTAAAGTAWAAATACMARCAGCCTGCAT TGC	1148
L7:25	TGTTTCCCTTTCTTCTTTAGCGACTTGAAGCTCTTGATC CAA	1149
L7:26	CAGAVCTTGTCGAGCTAATGGCGGCATACTATCAGTAGT AGG	1150
L7:27	CAGAVCTTGTCGGASTAATGGCGGCATACTATCAGTAGT AGG	1151
L7:28	GAATAAGAAGGCTGGCTCTGCACCTTGGTGATCCTTTAA TTC	1152
L7:29	TTTAAGTTGTTTTTCTAATCCGCATATGATCAATTCAAG GCC	1153
L7:30	GGGAACTTCGGCGCGCCTTAAGACCCACTTTCACA	1154

After transformation of the library into *E. coli* KM3 and plating on LB+Cb+Km the resulting colonies were reformatted into fifty-two 384-well microtiter plates (~20,000 colonies) and subsequently used to replica plate onto M9 assay medium containing either 0 μg/ml, 0.02 μg/ml or 0.2 μg/ml ethametsulfuron. After incubation at 30° C. for 48 hrs the plates were observed and 326 hits responding to 0.02 μg/ml inducer were identified and re-arrayed into 96-well format. Following incubation at 30° C. for 15, 24, 48 and 120 hours, digital images of the plates were taken and the relative colony color information converted to numerical data. DNA sequence analysis was carried out in parallel, and the two data sets merged for calculation of sequence-activity relationships. Sequence data for the top ten clones as well as the summary of the sequence-activity relationships are shown in Table 16. The results of sequence-activity relationship study revealed preferences impacting both activation and background activities of the putative ethametsulfuron repressors (EsR's). For example, one significant finding from this library was that modification L55M greatly reduced background activity and thus enhanced levels of fold activation. As seen from libraries L4 and L6, K108Q and wt Q109 were preferred for activation. There was also a high degree of bias towards L170V related to activation. This is different from the L170A or L170G bias seen in library L6, as those modifications had a strong correlation with lowering background activity in library L6. Finally, having a less dramatic effect on activation but nevertheless preferred is F67Y.

L. Induction Properties of Top L7 Hits in Liquid Cultures

Based on the performance of the re-arrayed hits, a second re-array was done and the clones tested for β-galactosidase activity in liquid culture along with wt TetR controls and selected 2^{nd} round hits to further analyze their performance and shuffling progress. Top hits from the L7 library were re-arrayed and tested in 96-well culture format for relative induction by 0.02 and 0.2 ug/ml inducer or background activity without inducer (FIG. **5**). Cultures were grown overnight and then subcultured into fresh medium containing appropriate treatments. Following six hours of incubation the cells were processed for enzyme assay. For assay of induced activity 5 μl of perforated cell mixture was used, for background activity 25 μl of cells were used such that detectable activity could be measured in the same time frame for all treatments. Background activity values were multiplied by ten to bring them into the range of the graph. The numbers below each sample refers to the final re-array well ID (vertical writing) and original re-array well ID (horizontal writing). The latter part of the graph contains the controls. Shown are 2nd round hits L4-89 and L4-120 as well as wt TetR. The final sample shows a control comprising wt TetR with 0.4 μM atc as cognate inducer for comparison. Results show that ten to fifteen of the top hits have induced activity approaching that of wt TetR induced with 0.4 μM atc. In addition, many of the hits have background activities nearly as low as wild type TetR. Some of the best hits have induction ratios with 0.2 ppm inducer (0.5 μM) approaching 70-80% of that of wt TetR (~1200-fold). It is interesting that the hits performing best at the low inducer concentration of 0.02 ppm (50 nM) also tended to have the higher background activity indicating that they are less tightly bound to tet operator and more easily released with transient inducer binding.

Comparison of induction activity between the 2^{nd} and 3^{rd} round hits is striking, showing greater than 200-fold improvement. Considering this improvement, a single additional round of shuffling and screening may yield sulfonylurea repressors (SuRs) that are nearly as sensitive to ethametsulfuron as the wt TetR is for tetracycline.

Summary

FIG. **9** provides a cumulative summary of the introduced diversity and observed amino acids in active SuRs obtained from the screening assays. Even though some positions were strongly biased i.e., observed more frequently in the selected population, as indicated by larger bolded type, the entirety of introduced diversity was observed in the full hit populations.

M. Novel Diversity Through In Vitro Mutagenesis

Residues A64, M86, C100, G104, F105, Q108, A113, S116, M134, Q135, I139, Y140, L147, L151, V170, L174, and K177 of round 3 hit L7-A11 were each mutagenized to all possible 20 amino acids to generate a set of 340 clones. Each of the clones was replica plated onto M9 assay medium containing 0, 5, 20 and 200 ppb ethametsulfuron. To assess relative activity of each of the mutants the plates containing ligand were photographed following 18 hrs of incubation at 37° C. To determine leakiness of the repressor clones the plate having no ligand addition was photographed after incubation for 24 hrs at 37° C. followed by 48 hrs of incubation at 25° C. Quantitative measurements were made by scanning digital photographs of each colony for blue color using ImageJ software.

These data revealed that select substitutions at positions L60, A64, N82, M86, A113, S116, M134, L174, and K177 demonstrated an increase in ethametsulfuron sensitivity relative to the parent clone L7-A11.

N. Fifth-Round Shuffling

Shuffling designed for improved ethametsulfuron sensitivity was performed. Library L13 (Table 18) was designed to incorporate novel diversity generated by the in vitro mutagenesis experiment in Example 1M that had either positive or neutral effect on activity. In addition, the library also incorporated diversity at selected cysteine residues in the backbone as listed (Table 18). The predicted library size is 124,000 members.

TABLE 18

	Library L13														
Residue	60	64	68	82	86	88	100	113	116	121	134	174	177	195	203
L7-A11	L	A	C	N	M	C	C	A	S	C	M	L	K	C	C
Diversity	L F	A D K	L C	K R N	M R	N	C A	A M	S C W	T	M F G	L I F Y	K H R	C S	C A

The library was assembled from synthetic oligonucleotides listed in Table 19 using methodology as described previously in this example.

TABLE 19

Oligo	Sequence	SEQ ID
L13:1	TGGCACGTCAAGAACAAGCGAGCTCTGCTAGACGCTATGGCC	1159
L13:2	ATTGAGATGTTSGATAGGCACAAGACCCACTACTGTCCTTTG	1160
L13:3	ATTGAGATGTTSGATAGGCACAAGACCCACTACCTGCCTTTG	1161
L13:4	ATTGAGATGTTSGATAGGCACGMCACCCACTACTGTCCTTTG	1162
L13:5	ATTGAGATGTTSGATAGGCACGMCACCCACTACCTGCCTTTG	1163
L13:6	GAAGGGGAAAGCTGGCAAGACTTCTTGAGGAACAATGCTAAG	1164
L13:7	GAAGGGGAAAGCTGGCAAGACTTCTTGAGGAACARGGCTAAG	1165
L13:8	TCCAKGAGAAATGCTTTGCTCAGTCACCGTGATGGAGCCAAG	1166
L13:9	GTCTGCCTAGGTACGGGCTTCACGGAGCAACAGTATGAAACT	1167
L13:10	GTCGCTCTAGGTACGGGCTTCACGGAGCAACAGTATGAAACT	1168
L13:11	GCTGAGAACTSKCTTGCCTTCCTGACACAACAAGGTTTCTCC	1169
L13:12	ATGGAGAACTSKCTTGCCTTCCTGACACAACAAGGTTTCTCC	1170
L13:13	CTTGAGAACGCCCTCTACGCATTTCAAGCTGTTGGGATCTAC	1171
L13:14	CTTGAGAACGCCCTCTACGCAGGTCAAGCTGTTGGGACTCTA	1172
L13:15	CTTGAGAACGCCCTCTACGCAATGCAAGCTGTTGGGATCTAC	1173
L13:16	ACTCTGGGTTGCGTCTTGCTGGATCAAGAGCTGCAAGTCGCT	1174
L13:17	AAGGAGGAGAGGGAAACACCTACTACTGATAGTATGCCGCCA	1175
L13:18	CTGGTTCGACAAGCTTACGAACTCCACGATCACCAAGGTGCA	1176
L13:19	CTGGTTCGACAAGCTTACGAACTCARAGATCACCAAGGTGCA	1177

TABLE 19-continued

Oligo	Sequence	SEQ ID
L13:20	CTGGTTCGACAAGCTHTCGAACTCCACGATCACCAAGGTGCA	1178
L13:21	CTGGTTCGACAAGCTHTCGAACTCARAGATGTGCACACCAAG	1179
L13:22	GAGCCAGCCTTCCTGTTCGGCCTTGAACTGATCATAWGTGGA	1180
L13:23	TTGGAGAAGCAGCTGAAGTGTGAAAGTGGGTCTTAATGATAG	1181
L13:24	TTGGAGAAGCAGCTGAAGGCAGAAAGTGGGTCTTAATGATAG	1182
L13:25	GTGCCTATCSAACATCTCAATGGCCATAGCGTCTAGCAGAGC	1183
L13:26	GTCTTGCCAGCTTTCCCCTTCCAAAGGACAGTAGTGGGTCTT	1184
L13:27	GTCTTGCCAGCTTTCCCCTTCCAAAGGCAGGTAGTGGGTCTT	1185
L13:28	GTCTTGCCAGCTTTCCCCTTCCAAAGGACAGTAGTGGGTGKC	1186
L13:29	GTCTTGCCAGCTTTCCCCTTCCAAAGGCAGGTAGTGGGTGKC	1187
L13:30	GAGCAAAGCATTTCTCMTGGACTTAGCATTGTTCCTCAAGAA	1188
L13:31	GAGCAAAGCATTTCTCMTGGACTTAGCCYTGTTCCTCAAGAA	1189
L13:32	GAAGCCCGTACCTAGGCAGACCTTGGCTCCATCACGGTGACT	1190
L13:33	GAAGCCCGTACCTAGAGCGACCTTGGCTCCATCACGGTGACT	1191
L13:34	GAAGGCAAGMSAGTTCTCAGCAGTTTCATACTGTTGCTCCGT	1192
L13:35	GAAGGCAAGMSAGTTCTCCATAGTTTCATACTGTTGCTCCGT	1193
L13:36	TGCGTAGAGGGCGTTCTCAAGGGAGAAACCTTGTTGTGTCAG	1194
L13:37	CAGCAAGACGCAACCCAGAGTGTAGATCCCAACAGCTTGAAA	1195
L13:38	CAGCAAGACGCAACCCAGAGTGTAGATCCCAACAGCTTGACC	1196
L13:39	CAGCAAGACGCAACCCAGAGTGTAGATCCCAACAGCTTGCAT	1197

TABLE 19-continued

Oligo	Sequence	SEQ ID
L13:40	AGGTGTTTCCCTCTCCTCCTTAGCGACTTGCAGCTCT TGATC	1198
L13:41	TTCGTAAGCTTGTCGAACCAGTGGCGGCATACTATCA GTAGT	1199
L13:42	TTCGADAGCTTGTCGAACCAGTGGCGGCATACTATCA GTAGT	1200
L13:43	GCCGAACAGGAAGGCTGGCTCTGCACCTTGGTGATCG TGGAG	1201
L13:44	GCCGAACAGGAAGGCTGGCTCTGCACCTTGGTGATCT YTGAG	1202
L13:45	ACACTTCAGCTGCTTCTCCAATCCACWTATGATCAGT TCAAG	1203
L13:46	TGCCTTCAGCTGCTTCTCCAATCCACWTATGATCAGT TCAAG	1204
L13:47	GCGCCAAGGTACCTTCTGCAGCTATCATTAAGACCCA CTTTC	1205

The assembled library was then cloned into pVER7571. This vector is the same as vector pVER7314 except for having a mutated ribosome binding site to reduce the amount of repressor produced per cell. This modification allows for more stringent assessment of repressor activity in the standard blue/white genetic plate assay, as well as in the liquid based whole cell quantitative β-galactosidase assay. Following plating of the library, approximately 5,000 clones were re-arrayed and replica plated onto M9 assay plates with no addition, or with 2 ppb ethametsulfuron plus 0.002% arabinose. Colonies responding the strongest with ethametsulfuron while remaining white without inducer were chosen as hits. One of the hits, L13-23, was found to be ~3-fold improved over the round 3 parent L7-A11 and to have the best repressor activity within this comparison (FIG. **11**). Sequence changes of the round 5 hit compared to parent molecule L7-A11 and wt TetR are shown in Table 20.

TABLE 20

Clone	55	60	64	67	68	82	86	88	100	104	105	108	113	116	121	134	135	139	140	147	151	170	174	177	195	203
wt	L	L	H	F	C	N	F	C	H	R	P	K	L	Q	C	L	S	H	F	E	H	L	I	F	C	C
L7-A11	M	–	A	Y	–	–	M	–	C	G	F	Q	A	S	–	M	Q	I	Y	L	L	V	L	K	–	–
L13-1-09	M	–	A	Y	–	K	M	N	A	G	F	Q	M	S	T	F	Q	I	Y	L	L	V	Y	K	–	A
L13-2-23	M	F	A	Y	–	K	M	N	A	G	F	Q	A	C	T	F	Q	I	Y	L	L	V	Y	K	–	–
L13-2-24	M	–	K	Y	L	–	M	N	C	G	F	Q	A	W	T	F	Q	I	Y	L	L	V	L	H	S	A

activity is reduced by ~90% in the presence of the effector and then derepressed following treatment with inducer.

Two ethametsulfuron repressors, EsR A11 and EsR D01, were selected for testing in conjunction with a wild type TetR control for dose response to ethametsulfuron by transient expression in *Nicotiana benthamiana* leaves (FIG. **6**). To this end, three test strains were derived by transformation of *Agrobacterium tumefaciens* EHA105 with three different T-DNA based vectors. *Agrobacterium* strains harboring binary vectors with a 35S::tetO-*Renilla* Luciferase reporter and dPCSV-tetR or -SuR effector variants were constructed. In addition to these tester cultures, an existing *Agrobacterium* strain harboring a dMMV-GFP T-DNA was added to the assay mixture to monitor the progression of *Agrobacterium* infection for sampling purposes.

To test the system for chemical switch activation, mixtures of tester *Agrobacterium* cultures containing 10% 35S::tetO-ReLuc reporter Agro, 10% dMMV-GFP Agro and 80% dPCSV-wt tetR Agro were infiltrated into *N. benthamiana* leaves and co-cultivated for 36 hours in the growth chamber. At this time the infiltrated leaves were excised and the petiole placed into water (negative control) or inducer at the test concentrations indicated in FIG. **6** and allowed to co-cultivate for another 36 hours. Infected leaf areas were assayed for *Renilla* luciferase activity and inducer treatments compared. The results show significant repression of reporter activity (~90%) with no inducer treatment (water control) for all tested repressors, and significant but incomplete induction of the EsR D01 repressor at inducer concentration as low as 0.02 ppm Es. Both EsR's were fully induced at 0.2 ppm Es whereas TetR was only fully induced at 2.0 ppm anhydrotetracycline (atc) (FIG. **6**).

B. High Throughput in Planta Assay Development Using *N. tabacum* BY-2 Cell Culture

In addition to the leaf assay it was desired to have an in planta assay to enable high throughput screening of SuR libraries for optimal plant functionality. We designed a system similar to the leaf assay but using tobacco BY-2 cell culture in 96-well format. BY-2 cell culture was transformed with a dMMV-HRA construct such that the culture would withstand treatment with target sulfonylurea test compounds. The resultant cell line grows and is fully resistant to 200 ppb chlorsulfuron.

Example 2

Plant Assay Development

A. *Nicotiana benthamiana* Leaf Infiltration Assay:

An in planta transient assay system was desired to rapidly confirm functionality of candidate SU-responsive repressors in planta prior to testing in transgenic plants. Therefore, an *Agrobacterium* based leaf infiltration assay was developed to measure repression and derepression activities. The strategy employed is to infiltrate leaves with a mixture of reporter and effector (repressor) *Agrobacterium* strains such that reporter

Example 3

Operator Binding Assay

To confirm that sulfonylurea ligands were binding directly to the modified repressor molecules and causing derepression, an in vitro tet operator gel shift study was undertaken.

An electrophoretic gel mobility shift assay (EMSA) of EsR variants was done to monitor binding to the tet operator (tetO)

sequence and response of the complex to inducers Es and Cs. TetO consists of a synthetic 48 bp tetO-containing fragment created from hybridization of oligonucleotide tetO1 (SEQ ID NO: 1155):

5'-CCTAATTTTTGTTGAC**ACTCTATCATTGATAGAGT**TATTTTACCAC

TC-3'

and complementary oligonucleotide tetO2 (SEQ ID NO: 1156):

5'-GGATTAAAAACAACTG**TGAGATAGTAACTATCTCA**ATAAAATGGTG

AG-3'

The tet operator is shown in bold.

An oligonucleotide and its complement of the same size containing no palindromic sequence was used as a control (SEQ ID NO: 1157):

5'-CCTAATTTTTGTTGACTGTGTTAGTCCATAGCTGGTATTTTACCAC

TC-3'

and complementary oligonucleotide (SEQ ID NO: 1158):

5'-GGATTAAAAACAACTGACACAATCAGGTATCGACCATAAAATGGTG

AG-3'

Five pmol of TetO or control DNA was mixed with the indicated amounts (FIG. **7**) of ethametsulfuron repressor protein (L7A11) or BSA control with or without inducer in complex buffer containing 20 mM Tris-HCl (pH8.0) and 10 mM EDTA. The mixture was incubated at room temperature for 0.5 hour before loading onto the gel. The reaction was electrophoresed on a Novex 6% DNA retardation gel (Invitrogen, EC6365BOX) at room temperature, 38 V in 0.5×TBE buffer for about 2 hours. DNA was detected by ethidium bromide staining. The DNA size marker consists of the low DNA mass ladder (InVitrogen 10068-013).

The results are shown in FIG. **7**. These results directly demonstrate that the modified repressors bind to operator DNA (lane 1 vs. lanes 3-5) and then are released from the operator sequence in an inducer-specific and dose dependent manner. The data also indicate an inducer preference for operator release by Es compared to Cs (lane 9 vs. 10). No change in operator release could be detected by atc compared to no inducer (lane 5 vs. 11).

Example 4

Binding and Dissociation Constants

Select SU repressors were further characterized for operator and ligand binding, affinity and dissociation kinetics using Biacore™ SPR technology (Biacore, GE Healthcare, USA). The technology is based on surface plasmon resonance (SPR), an optical phenomenon that enables detection of unlabeled interactants in real time. The SPR-based biosensors can be used in determination of active concentration, screening and characterization in terms of both affinity and kinetics.

The kinetics of an interaction, i.e., the rates of complex formation (k_a) and dissociation (k_d), can be determined from the information in a sensorgram. If binding occurs as sample passes over a prepared sensor surface, the response in the sensorgram increases. If equilibrium is reached, a constant signal is seen. Replacing the sample with buffer causes the bound molecules to dissociate and the response decreases. Biacore evaluation software generates the values of k_a and k_d by fitting the data to interaction models.

The affinity of an interaction is determined from the level of binding at equilibrium (seen as a constant signal) as a function of sample concentration. Affinity can also be determined from kinetic measurements. For a simple 1:1 interaction, the equilibrium constant K_D is the ratio of the kinetic rate constants, k_d/k_a.

A. Operator Binding Characterization of Repressors

Repressor	k_a ($M^{-1} s^{-1}$)	K_d (s^{-1})	K_D (nM)
Wt TetR	3.3×10^5	3.0×10^{-3}	9.0 ± 1.0
L7-1C03-A5	4.7×10^4	7.8×10^{-3}	150 ± 5
L7-3E03-D1	5.5×10^4	1.1×10^{-2}	200 ± 50
L7-1F08-A11	7.1×10^4	1.7×10^{-2}	250 ± 120
L7-1G06-B2	4.6×10^4	1.9×10^{-2}	430 ± 160

B. SU Binding Characterization of Repressors

	KD (M)				
Repressor	Es + Mg	Es – Mg	Cs + Mg	Cs – Mg	ATC + Mg
L7-1C03-A5	0.46	1.78	83	365	Null
L7-1F08-A11	0.45	1.09	40	92	Null
L7-1G06-B2	0.53	2.15	60	255	Null
L7-3E03-D1	0.73	2.15	48	115	Null
Wt TetR	Null	Null	Null	Null	0.0036

Example 5

Sulfonyl Urea Repressor Ligand-Binding Domain Fusions

The ligand binding domains from the sulfonylurea repressors provided herein can be fused to alternative DNA binding domains in order to create further sulfonylurea repressors that selectively and specifically bind to other DNA sequences (e.g., Wharton and Ptashne (1985) *Nature* 316:601-605). Many domain swapping experiments have been published, demonstrating the breadth and flexibility of this approach. Generally, an operator binding domain or specific amino acid/operator contact residues from a different repressor system will be used, but other DNA binding domains can also be used. For example, a polynucleotide encoding a TetR(D)/SuR chimeric polypeptide consisting of the DNA binding domain from TetR(D) (e.g., amino acid residues 1-50) and ligand binding domain of a SuR residues (e.g., amino acid residues 51-208 from TetR(B) can be constructed using any standard molecular biology method or combination thereof, including restriction enzyme digestion and ligation, PCR, synthetic oligonucleotides, mutagenesis or recombinational cloning. For example, a polynucleotide encoding a SuR comprising a TetR(D)/SuR chimera can be constructed by PCR (Landt et al. (1990) *Gene* 96:125-128; Schnappinger et al. (1998) *EMBO J* 17:535-543) and cloned into a suitable expression cassette and vector. Any other TetOp binding domains can be substituted to produce a SuR that specifically binds to the cognate tet operator sequence.

In addition, mutant TetO^c binding domains from variant TetR's having suppressor activity on constitutive operator sequences (tetO-4C and tetO-6C) can be used (see, e.g., Helbl and Hillen (1998) *J Mol Biol* 276:313-318; and Helbl et al. (1998) *J Mol Biol* 276:319-324). Further, the polynucleotides encoding these DNA binding domains can be modified to change their operator binding properties. For example, the polynucleotides can be shuffled to enhance the binding strength or specificity to a wild type or modified tet operator sequence, or to select for specific binding to a new operator sequence.

Additional variants could be made by fusing an SuR repressor, or an SuR ligand binding domain to an activation domain. Such systems have been developed using Tet repressors. For example, one system converted a tet repressor to an activator via fusion of the repressor to a transcriptional transactivation domain such as herpes simplex virus VP16 and the tet repressor (tTA, Gossen and Bujard (1992) Proc Natl Acad Sci USA 89:5547-5551). The repressor fusion is used in conjunction with a minimal promoter which is activated in the absence of tetracycline by binding of tTA to tet operator sequences. Tetracycline inactivates the transactivator and inhibits transcription.

Example 6

Testing of Repressor Proteins in Soybean

Any transformation protocols, culture techniques, soybean source, and media, and molecular cloning techniques can be used with the compositions and methods.

A Transformation and Regeneration of Soybean (*Glycine max*)

Transgenic soybean lines are generated by the method of particle gun bombardment (Klein et al. Nature 327:70-73 (1987); U.S. Pat. No. 4,945,050) using a BIORAD Biolistic PDS1000/He instrument and either plasmid or fragment DNA. The following stock solutions and media are used for transformation and regeneration of soybean plants:

Stock Solutions:

Sulfate 100× Stock: 37.0 g MgSO4.7H2O, 1.69 g MnSO4.H2O, 0.86 g ZnSO4.7H2O, 0.0025 g CuSO4.5H2O

Halides 100× Stock: 30.0 g CaCl2.2H2O, 0.083 g KI, 0.0025 g CoCl2.6H2O

P, B, Mo 100× Stock: 18.5 g KH2PO4, 0.62 g H3BO3, 0.025 g Na2MoO4.2H2O

Fe EDTA 100× Stock: 3.724 g Na2EDTA, 2.784 g FeSO4.7H2O

2,4-D Stock: 10 mg/mL 2,4-Dichlorophenoxyacetic acid

B5 vitamins, 1000× Stock: 100.0 g myo-inositol, 1.0 g nicotinic acid, 1.0 g pyridoxine HCl, 10 g thiamine.HCL.

Media (Per Liter):

SB199 Solid Medium: 1 package MS salts (Gibco/BRL, Cat. No. 11117-066), 1 mL B5 vitamins 1000× stock, 30 g Sucrose, 4 ml 2,4-D (40 mg/L final concentration), pH 7.0, 2 g Gelrite

SB1 Solid Medium: 1 package MS salts (Gibco/BRL, Cat. No. 11117-066), 1 mL B5 vitamins 1000× stock, 31.5 g Glucose, 2 mL 2,4-D (20 mg/L final concentration), pH 5.7, 8 g TC agar

SB196: 10 mL of each of the above stock solutions 1-4, 1 mL B5 Vitamin stock, 0.463 g (NH4)2SO4, 2.83 g KNO3, 1 mL 2,4 D stock, 1 g asparagine, 10 g sucrose, pH 5.7

SB71-4: Gamborg's B5 salts, 20 g sucrose, 5 g TC agar, pH 5.7.

SB103: 1 pk. Murashige & Skoog salts mixture, 1 mL B5 Vitamin stock, 750 mg MgCl2 hexahydrate, 60 g maltose, 2 g gelrite, pH 5.7.

SB166: SB103 supplemented with 5 g per liter activated charcoal.

Soybean Embryogenic Suspension Culture Initiation:

Soybean cultures are initiated twice each month with 5-7 days between each initiation. Pods with immature seeds from available soybean plants 45-55 days after planting are picked, removed from their shells and placed into a sterilized magenta box. The soybean seeds are sterilized by shaking them for 15 min in a 5% Clorox solution with 1 drop of ivory soap (i.e., 95 mL of autoclaved distilled water plus 5 mL Clorox and 1 drop of soap, mixed well). Seeds are rinsed using 2 1-liter bottles of sterile distilled water and those less than 3 mm are placed on individual microscope slides. The small end of the seed is cut and the cotyledons pressed out of the seed coat. Cotyledons are transferred to plates containing SB199 medium (25-30 cotyledons per plate) for 2 weeks, then transferred to SB1 for 2-4 weeks. Plates are wrapped with fiber tape. After this time, secondary embryos are cut and placed into SB196 liquid media for 7 days.

Culture Conditions:

Soybean embryogenic suspension cultures (cv. Jack) are maintained in 50 mL liquid medium SB196 on a rotary shaker, 150 rpm, 26° C. with cool white fluorescent lights on 16:8 h day/night photoperiod at light intensity of 60-85 μE/m2/s. Cultures are subcultured every 7 days to two weeks by inoculating approximately 35 mg of tissue into 50 mL of fresh liquid SB196 (the preferred subculture interval is every 7 days).

Preparation of DNA for Bombardment:

In particle gun bombardment procedures it is possible to use purified intact plasmid DNA; or DNA fragments containing only the recombinant DNA expression cassette(s) of interest. For every seventeen bombardment transformations, 85 μL of suspension is prepared containing 1 to 90 picograms (pg) of plasmid DNA per base pair of each DNA plasmid. Both recombinant DNA plasmids are co-precipitated onto gold particles as follows. The DNAs in suspension are added to 50 μL of a 10-60 mg/mL 0.6 μm gold particle suspension and then combined with 50 μL CaCl2 (2.5 M) and 20 μL spermidine (0.1 M). The mixture is vortexed for 5 sec, spun in a microfuge for 5 sec, and the supernatant removed. The DNA coated particles are then washed once with 150 μL of 100% ethanol, vortexed and pelleted, then resuspended in 85 μL of anhydrous ethanol. Five μL of the DNA coated gold particles are then loaded on each macrocarrier disk.

Tissue Preparation and Bombardment with DNA:

Approximately 150 to 250 mg of two-week-old suspension culture is placed in an empty 60 mm×15 mm petri plate and the residual liquid removed from the tissue using a pipette. The tissue is placed about 3.5 inches away from the retaining screen and each plate of tissue is bombarded once. Membrane rupture pressure is set at 650 psi and the chamber is evacuated to −28 inches of Hg. Following bombardment, the tissue from each plate is divided between two flasks, placed back into liquid media, and cultured as described above.

Selection of Transformed Embryos and Plant Regeneration:

After bombardment, tissue from each bombarded plate is divided and placed into two flasks of SB196 liquid culture maintenance medium per plate of bombarded tissue. Seven days post bombardment, the liquid medium in each flask is replaced with fresh SB196 culture maintenance medium supplemented with 100 ng/ml selective agent (selection medium). Transformed soybean cells can be selected using a sulfonylurea (SU) compound such as 2 chloro N ((4 methoxy

6 methy 1,3,5 triazine 2 yl)aminocarbonyl)benzenesulfonamide (common names: DPX-W4189 and chlorsulfuron). Chlorsulfuron (Cs) is the active ingredient in the DuPont sulfonylurea herbicide, GLEAN®. The selection medium containing SU is replaced every two weeks for 6-8 weeks. After the 6-8 week selection period, islands of green, transformed tissue are observed growing from untransformed, necrotic embryogenic clusters. These putative transgenic events are isolated and kept in SB196 liquid medium with Cs at 100 ng/ml for another 2-6 weeks with media changes every 1-2 weeks to generate new, clonally propagated, transformed embryogenic suspension cultures. Embryos spend a total of around 8-12 weeks in contact with Cs. Suspension cultures are subcultured and maintained as clusters of immature embryos and also regenerated into whole plants by maturation and germination of individual somatic embryos.

Somatic embryos became suitable for germination after four weeks on maturation medium (1 week on SB166 followed by 3 weeks on SB103). They are then removed from the maturation medium and dried in empty petri dishes for up to seven days. The dried embryos are then planted in SB71 4 medium where they are allowed to germinate under the same light and temperature conditions as described above. Germinated embryos are transferred to potting medium and grown to maturity for seed production.

B. Vector Construction and Testing

Plasmids were made using standard procedures and from these plasmids DNA fragments were isolated using restriction endonucleases and agarose gel purification according to the protocol described in Example 6A. Each DNA fragment contained three cassettes. Cassette 1 is a reporter expression cassette; Cassette 2 is the repressor expression cassette; and, Cassette 3 is an expression cassette providing an HRA gene. The repressors tested in Cassette 2 are described in Table 21. The polynucleotides comprising the repressor coding region were synthesized to comprise plant preferred codons. In all cases Cassette 1 contained a 35S cauliflower mosaic virus promoter having three tet operators introduced near the TATA box (Gatz et al. (1992) *Plant J* 2:397-404 (3XOpT 35S)) driving expression of DsRed followed by the 35S cauliflower mosaic virus 3' terminator region. In all cases cassette three contained the S-adenosylmethionine synthase promoter followed by the HRA version of the acetolactase synthase (ALS) gene followed by the *Glycine max* ALS 3' terminator. The HRA version of the ALS gene confers resistance to sulfonylurea herbicides. EF1A1 is the promoter of a soybean translation elongation factor EF1 alpha described in patent publication US20080313776.

DNA fragments were used for soybean transformation according to the protocol described in Example 6A. Plants were regenerated and leaf discs (~0.5 cm) were harvested from young leaves. The leaf discs were incubated in SB103 liquid media containing 0 ppm, 0.5 ppm or 5 ppm ethametsulfuron for 2-5 days. Ethametsulfuron (product number PS-2183) was purchased from Chem Service (West Chester, Pa.) and solubilized in either 10 mM NaOH or 10 mM NH_4OH. On each day leaf discs were examined under a dissecting microscope with a DsRed band pass filter. The presence of DsRed was scored visually.

Plants that expressed DsRed at 0 time were scored as leaky. Plants that did not express DsRed after five days were scored as negative. Plants that expressed DsRed after addition of ethametsulfuron were scored as inducible. Results from plants derived from DNAs described in Table 21 are shown in Table 22.

TABLE 22

Name	Alias	Total Events	% Leaky	% Negative	% Inducible
PHP37586A	CHSW004	12	33	33	33
PHP37587A	CHSW005	28	7	50	43
PHP37588A	CHSW006	6	0	0	100
PHP37589A	CHSW007	9	0	22	78
PHP39389A	CHSW010*	19	5	26	42
PHP39390A	CHSW011*	35	0	17	57

*In these cases the total does not equal 100% as multiple plants were examined from some events and, in some cases, different plants from the same event behaved differently.

This shows that the repressor protein responds to ethametsulfuron by inducing expression of DsRed. Plants derived from the first four fragments showed visual evidence of DsRed after three days of incubation. Plants derived from the last two fragments showed visual evidence of DsRed after two days of incubation. The presence of DsRed was scored visually, but this was confirmed by Western Blot analysis on a selection of transformants using a rabbit polyclonal antibody (ab41336) from Abcam (Cambridge, Mass.).

Leaf punches were harvested as described above from a selection of transformants and incubated in SB103 media with 0, 5, 50, 250 and 500 ppb ethametsulfuron. At all concentrations of ethametsulfuron, leaves showed visual evidence of DsRed after three days of incubation. At the lowest concentration (5 ppb) the presence of DsRed was limited to a "halo" near the outside edge of the leaf disc.

Plants were allowed to mature as described in Example 6A. Since soybeans are self fertilizing, the T1 seeds derived from

TABLE 21

Fragment Name	Fragment alias	Cassette 2	Repressor alias	Repressor SEQ ID	Fragment SEQ ID
PHP37586A	CHSW004	EF1A1::EsR1::Nos3'	L7-IC3-A5	1240	1222
PHP37587A	CHSW005	EF1A1::EsR2::Nos3'	L7-1F8-A11	1241	1223
PHP37588A	CHSW006	EF1A1::EsR2::Nos3	L7-1G6-B2	1242	1224
PHP37589A	CHSW007	EF1A1::EsR4::Nos3'	L7-3E3-D1	1243	1225
PHP39389A	CHSW010	EF1A1::EsR5::CaMV35S3'	L12-1-10	1232	1226
PHP39390A	CHSW011	EF1A1::EsR6::CaMV35S3'	L13-2-23	1233	1227

these plants would be expected to segregate 1 wild type: 2 hemizygote: 1 homozygote if only one transgene locus was created during transformation. Sixteen seeds from five different events were planted and allowed to germinate. Leaf punches were collected from young seedlings and incubated in SB103 media with 0 and 5 ppm ethametsulfuron. Leaf discs were scored for DsRed expression and 0 and 3 days and results are shown in Table 23.

TABLE 23

Name	Event ID	Total # Seeds Germinated	# Leaky Plants	# Negative Plants	# Inducible Plants
PHP37586A	6048.3.8.3	11	0	2	9
PHP37587A	6049.2.2.4	12	0	5	7
PHP37588A	6150.3.2.1	14	0	1	13
PHP37589A	6154.4.5.1	15	0	15	0
PHP39389A	6151.4.18.1	12	3	9	0

Example 7

Testing of Repressor Proteins in Corn

To evaluate SU repressors in plants, RFP reporter constructs were constructed and transformed into maize cells via *Agrobacterium* using the following T-DNA configuration: RB-35S/TripleOp/Pro::RFP-Ubi Pro::EsR-HRA cassette-PAT cassette-LB.

Using standard molecular biology and cloning techniques, T-DNA vectors having the configuration above comprising selected round 3 SU repressors (EsRs) were constructed. The polynucleotides comprising the repressor coding region were synthesized to comprise plant preferred codons. The constructs are summarized below:

Construct ID	SU repressor alias (EsR)	SU repressor SEQ ID
PHP37707	L7-1C3-A5	1240
PHP37708	L7-1F8-A11	1241
PHP37709	L7-1G6-B2	1242
PHP37710	L7-3E3-D1	1243

The reporter construct T-DNA contained a 35S promoter with two tet operators flanking the TATA box and one downstream adjacent to the transcription start site (as described by Gatz et al. (1992) Plant Cell 2:397-404) driving expression of the Red Fluorescent Protein gene, a ubiquitin driven SU repressor (EsR), an expression cassette containing the maize HRA gene for SU resistance and a moPAT expression cassette for selection.

Immature embryos were transformed using standard methods and media. Briefly, immature embryos were isolated from maize and contacted with a suspension of *Agrobacterium*, to transfer the T-DNA's containing the sulfonylurea expression cassette to at least one cell of at least one of the immature embryos. The immature embryos were immersed in an *Agrobacterium* suspension for the initiation of inoculation and cultured for seven days. The embryos were then transferred to culture medium containing carbinicillin to kill off any remaining *Agrobacterium*. Next, inoculated embryos were cultured on medium containing both carbinicillin and bialaphos (a selective agent) and growing transformed callus was recovered. The callus was then regenerated into plantlets on solid media before transferring to soil to produce mature plants. Approximately 10 single copy events from each of the constructs were sent to the greenhouse.

To evaluate de-repression, callus was transferred to medium with and without ethametsulfuron and chlorsulfuron and RFP Fluorescence was observed under the microscope (see FIG. **10**A). Most events de-repressed and there were no obvious differences between the round three repressors tested. To evaluate de-repression in plants, seeds for single copy plants were germinated in the presence of ethamethsulfron and fluorescence was observed and photographed (see FIG. **10**B). As a positive control, a vector containing the same configuration of expression cassettes as PHP37707-10, but with UBI::TetR in place of UBI::EsR, were transformed into maize immature embryos and tested for induction on doxycline. When grown in the presence of 1 mg/l doxycycline, transgenic callus and plants containing the TetR expression cassette induced over a similar 5-6 day period.

The articles "a" and "an" refer to one or more than one of the grammatical object of the article. By way of example, "an element" means one or more of the element. All book, journal, patent publications and grants mentioned in the specification are indicative of the level of those skilled in the art. All publications and patent applications are herein incorporated by reference to the same extent as if each individual publication or patent application was specifically and individually indicated to be incorporated by reference. Although the foregoing invention has been described in some detail by way of illustration and example for purposes of clarity of understanding, certain changes and modifications may be practiced within the scope of the appended claims. These examples and descriptions are illustrative and are not read as limiting the scope of the appended claims.

SEQUENCE LISTING

The patent contains a lengthy "Sequence Listing" section. A copy of the "Sequence Listing" is available in electronic form from the USPTO web site (http://seqdata.uspto.gov/?pageRequest=docDetail&DocID=US08580556B2). An electronic copy of the "Sequence Listing" will also be available from the USPTO upon request and payment of the fee set forth in 37 CFR 1.19(b)(3).

What is claimed:

1. An isolated polypeptide comprising a sulfonylurea-responsive repressor that specifically binds to a polynucleotide comprising an operator sequence, wherein the binding is regulated by a sulfonylurea compound, and wherein said polypeptide comprises an amino acid sequence having at least 90% sequence identity to SEQ ID NO:1233.

2. The polypeptide of claim **1**, wherein said polypeptide comprises an amino acid sequence having at least 95% sequence identity to SEQ ID NO:1233.

* * * * *